



SmallMap: Low-cost Community Road Map Sensing with Uncertain Delivery Behavior

ZHIQING HONG, JD Logistics, China and Rutgers University, USA

HAOTIAN WANG, JD Logistics, China

YI DING, University of Texas at Dallas, USA

GUANG WANG*, Florida State University, USA

TIAN HE, JD Logistics, China

DESHENG ZHANG, Rutgers University, USA

Accurate road networks play a crucial role in modern mobile applications such as navigation and last-mile delivery. Most existing studies primarily focus on generating road networks in open areas like main roads and avenues, but little attention has been given to the generation of community road networks in closed areas such as residential areas, which becomes more and more significant due to the growing demand for door-to-door services such as food delivery. This lack of research is primarily attributed to challenges related to sensing data availability and quality. In this paper, we design a novel framework called SmallMap that leverages ubiquitous multi-modal sensing data from last-mile delivery to automatically generate community road networks with low costs. Our SmallMap consists of two key modules: (1) a *Trajectory of Interest Detection module* enhanced by exploiting multi-modal sensing data collected from the delivery process; and (2) a *Dual Spatio-temporal Generative Adversarial Network module* that incorporates Trajectory of Interest by unsupervised road network adaptation to generate road networks automatically. To evaluate the effectiveness of SmallMap, we utilize a two-month dataset from one of the largest logistics companies in China. The extensive evaluation results demonstrate that our framework significantly outperforms state-of-the-art baselines, achieving a precision of 90.5%, a recall of 87.5%, and an F1-score of 88.9%, respectively. Moreover, we conduct three case studies in Beijing City for courier workload estimation, Estimated Time of Arrival (ETA) in last-mile delivery, and fine-grained order assignment.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**.

Additional Key Words and Phrases: Community road network, Crowdsensing, Mobile sensing, Last-mile delivery

ACM Reference Format:

Zhiqing Hong, Haotian Wang, Yi Ding, Guang Wang, Tian He, and Desheng Zhang. 2024. SmallMap: Low-cost Community Road Map Sensing with Uncertain Delivery Behavior. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 2, Article 50 (June 2024), 26 pages. <https://doi.org/10.1145/3659596>

*Dr. Guang Wang is the corresponding author.

Authors' Contact Information: [Zhiqing Hong](#), JD Logistics, China and Rutgers University, USA, zhiqing.hong@rutgers.edu; [Haotian Wang](#), JD Logistics, China, wanghaotian18@jd.com; [Yi Ding](#), University of Texas at Dallas, USA, yi.ding@utdallas.edu; [Guang Wang](#), Florida State University, USA, guang@cs.fsu.edu; [Tian He](#), JD Logistics, China, tim.he@jd.com; [Desheng Zhang](#), Rutgers University, USA, desheng@cs.rutgers.edu.

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, to 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 2474-9567/2024/6-ART50

<https://doi.org/10.1145/3659596>



Fig. 1. Motivation. (a) Community roads and open roads. (b) Commercial maps with missing community road networks. (c) Failure of travel time estimation without community road networks.

1 INTRODUCTION

Accurate and fine-grained community road networks, roads in closed or semi-closed areas such as residential neighborhoods and industrial parks that typically have lower speed limits and may have restrictions on vehicle access, are important for emerging door-to-door services [64, 65] (e.g., last-mile delivery and on-demand food delivery) because these service workers (e.g., couriers) rely on searching for accurate community locations to fulfill tasks, e.g., travel time estimation. However, road networks in many communities (especially those newly built) are not captured by current commercial navigation software (e.g., Google and Baidu Maps [1]) and open-source geographic database [5]) (as in Fig. 1(a) and Fig. 1(b)), which brings huge challenges for efficient door-to-door services. For example, as in Fig. 1(c) the service provider cannot accurately estimate the delivery time without community road networks, which potentially decreases business efficiency and user satisfaction. Hence, in this work, we aim to design a novel framework to automatically generate community road networks.

To date, various approaches have been proposed for the generation of open road networks [10, 16, 23, 63, 70, 82], i.e., road networks designed for higher-speed vehicular traffic such as avenues and highways. These roads connect different urban and suburban areas with fewer access restrictions. These studies cannot be applied to community road network generation directly due to the following three reasons. (i) Most communities, especially residential areas in Chinese cities, are closed, i.e., no through traffic is allowed to get inside of communities. Therefore, existing studies utilizing public vehicles' trajectories (such as taxis) are not applicable [16, 76] for community road network generation. Also, in many communities, some restrictions prohibit the entry of public bikes and commercial data collection vehicles like Google Maps cars or Baidu Maps cars, for pedestrian safety and regulatory purposes. In addition, it brings a high cost to deploy a specific system like drones for collecting and updating community maps. (ii) The characteristics of community roads are different. The diverse shapes (e.g., circular, undulating, and straight) and narrow widths (e.g., 2-meter width) of community roads, combined with GPS inaccuracies [19], pose significant challenges to generating community roads. (iii) Roads in communities can be obscured by trees or buildings, severely limiting the effectiveness of image-based methods [10, 23, 70].

Even though some community road networks are included by commercial navigation companies like Google Maps [3] and Gaode Maps [2], it poses two practical challenges for delivery companies to directly use these existing data. Firstly, utilizing commercial community road networks could potentially leak critical business information such as order distribution and real-time courier locations, which is similar to the privacy risk in vehicle-based urban sensing [8]. Secondly, even though commercial maps are free for individual users, delivery companies need to pay a high cost for frequent usage [40]. Therefore, it is necessary and important to design a low-cost approach to generate community road networks.

Recent advancements in door-to-door delivery services present us with an unprecedented opportunity. Experienced delivery couriers regularly traverse in communities, thereby enabling the collection of rich multi-modal sensing data (e.g., GPS and magnetic field sensor data) using their Personal Digital Assistants (PDAs, which are similar to smartphones). The continuously collected sensing data has the potential to capture structural details of the road networks. However, through our data analysis, we found that utilizing this multi-modal sensing data for the community road network generation is nontrivial because of three main challenges. **(i) Uncertain courier behavior.** Couriers' mobility within communities is highly uncertain, including a mixture of indoor activities (such as package delivery), outdoor movements, and stationary trajectories, which makes it difficult to discern trajectories that fall within the road networks. **(ii) Inaccurate sensing.** Due to the hardware limitations and environmental effects, the GPS trajectories recorded from devices carried by couriers may contain uncertain measurement errors, ranging from a few meters to 100 meters, in complex environments [12, 19]. **(iii) Unbalanced sensing.** Despite a high penetration rate in communities, different road segments within a community are visited with varying frequencies. Details are introduced in Section 3.3.

To address the above challenges, in this paper, we design SmallMap a low-cost community road network generation framework based on multi-modal sensing data. Specifically, within the SmallMap framework: (i) We design a novel Trajectory-of-Interest Detection module to eliminate noisy trajectories caused by uncertain courier behaviors via discerning courier statuses from multi-modal sensing data. (ii) We combine aggregated trajectories and design a **Dual Spatio-Temporal Generative Adversarial Network (DualSTGAN)** to generate road networks that preserve real-world topology and geometry despite the noisy sampled trajectories. DualSTGAN is designed to model the shape information displayed by the aggregated trajectories, which consists of a RoadGAN module to generate road networks from aggregated trajectories that address the challenge of inaccurate sensing, and a LineGAN module to enrich the generated road structures with low revisit frequencies and address the challenge of unbalanced sensing. We summarize our contribution as follows.

- To our knowledge, we conduct the first study to generate community road networks by utilizing noisy sampled multi-modal sensing data from last-mile delivery. The generated fine-grained community road networks have great potential to benefit various emerging door-to-door services.
- Technically, we design a low-cost community map generation framework called SmallMap, which consists of two main modules, a Trajectory-of-Interest Detection module to address uncertain courier status challenge, and a Dual Spatio-temporal Generative Adversarial Network (DualSTGAN) module to address inaccurate sensing and unbalanced sensing challenges.
- More importantly, we collect a real-world dataset in two Chinese cities, including 12,336 trajectories and over 3.16 million GPS points to evaluate the performance of SmallMap quantitatively and qualitatively. Extensive evaluation results show that SmallMap achieves an overall F1-score of 88.9%. Furthermore, by working with the logistics company we are collaborating with, we have conducted case studies in the Chinese city of Beijing. The results of case studies demonstrate the utility of the generated road networks on three fundamental tasks in last-mile delivery.

2 RELATED WORK

2.1 Road Network Generation

With the growing importance of road networks in smart cities, road network generation (or road network inference) has attracted much interest in research communities [9, 10, 14, 17, 24, 27, 34, 36, 43, 50, 54, 58, 63, 63, 70, 83]. Most existing studies fall into three categories, i.e., inference based on GPS trajectories [13, 15, 21, 26, 69, 76, 86], inference based on aerial images [10, 23, 63, 70, 83], and inference based on the combination of trajectory and aerial images [63, 71, 77]. Zhang et al. [83] generate a map (houses and roads) using high-resolution aerial images by designing a U-Net-based deep learning model. Fang et al. [27] infer road network structure and categories by

modeling complementary trajectories such as taxi and truck. Another line of work combines different modalities of data such as images and trajectories [63, 71, 77], which explore the semantic information in images and the sequence information in trajectories. While most existing works focus on road networks in open areas for vehicles, Li et al. [50] utilize on-demand delivery data to generate community road networks. However, this work cannot be applied to our scenario directly due to the uncertain quality of trajectory data brought by uncertain courier status in last-mile delivery. Also, Cao et al. [14] infer walkways utilizing trajectories and step counts derived from mobile devices, which differs from our scenario where the detailed step counts are not available and the topology of road segments in communities are more complex.

2.2 Road Network and Trajectory Mining

The road network is closely tight to human mobility and thus attracts lots of research interests in building practical applications upon road networks, e.g., traffic prediction [51], map matching [68, 72], trajectory recovery [74], and localization [44, 78]. Spatio-temporal trajectories are generated every day with a huge volume and spatial coverage, which inspires much research in this direction, e.g., trajectory data management [31, 87], trajectory modeling and representation [25, 51, 56, 80, 81], and trajectory-based sensing [37, 38, 59, 73]. Based on vehicle trajectories, Li et al. [51] propose a deep learning model to predict urban traffic flow that can potentially benefit downstream traffic management. Due to the potentially wide coverage rate, trajectory enables many urban service applications. Based on crowdsourced bike trajectories, He et al. [38] design a framework to detect illegal parking vehicles along city roads, which utilizes the spatiotemporal trajectory patterns.

2.3 Urban Sensing

By harnessing multi-modal sensing data, there is a growing interest in the research community to measure the physical phenomenon for smart city applications, e.g., human status estimation [33, 88, 89], human mobility modeling [45, 79], traffic analysis [18, 61, 74, 85], last-mile delivery [41, 42], and urban boundary identification [20]. Isaacman et al. [45] design a model based on Call Detail Records (CDRs) [39] to model human mobility in multiple metropolitan areas and evaluate it on large-scale datasets. Based on open government data, Chen et al. [20] propose a framework to estimate urban village boundaries and population. Crowdsensing provides a great opportunity to obtain multi-modal sensing data on a large scale to support various urban sensing tasks, but few existing works utilize multi-modality crowdsensing data for community road network generation.

In summary, our work is different from existing studies because of the following aspects.

- (1) We focus on community road network generation, which is an extension of existing studies but with significant differences. GPS has more drifts in communities due to the impact of environmental factors such as high buildings and trees. Also, community roads are mainly designed for pedestrians and thereby are narrow and of casual shapes, which makes our problem more challenging.
- (2) We design a framework based on both online and offline inference, which reduces the storage and data transition costs significantly. Moreover, a DualSTGAN with a novel data augmentation approach is presented to generate high-quality road networks with low costs, which also benefits related studies to address the challenge of training data shortage.
- (3) Most existing studies utilize an existing map (such as OSM [5]) and complement existing maps with GPS trajectories. On the contrary, we focus on generating road networks from scratch, since many communities' roads have not been charted by existing open-sourced maps like OSM.

3 COURIER SENSING SYSTEM AND MOTIVATION

In this section, we introduce the design of the courier sensing system, multi-modal sensing data, opportunities, and challenges in generating community road networks with multi-modal courier sensing data.

3.1 Sensing System and Data Collection

In Fig. 2, we illustrate the overall process of last-mile delivery, which integrates three main processes, i.e., in the delivery station, on the road, and in communities for parcel delivery. (1) From t_0 to t_1 , the courier picks the parcels up when the parcels arrive at the delivery station. (2) From t_1 to t_2 , the courier carries the parcels and drives on open roads to customers' locations. (3) From t_2 to t_4 , the courier delivers parcels in communities. The courier carries PDAs during the delivery process, which can be utilized to collect heterogeneous sensing data that covers road networks in communities from different types of sensors. In this study, we especially focus on the duration from t_2 to t_4 , where couriers are in communities.

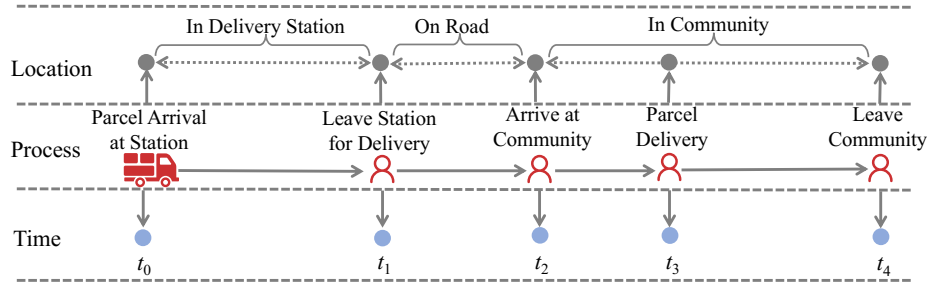


Fig. 2. Last-mile delivery illustration.

In last-mile delivery, the real-time location information of couriers is important for delivery time estimation and courier scheduling. Normally, a courier carries a PDA, starting from the delivery station to one or multiple communities for parcel delivery. Inside the community, couriers frequently change their statuses, moving between indoors and outdoors. For courier localization and sensing data collection, we design a localization module and deploy it at couriers' PDAs with various operating systems (from Android 4 to Android 11) and hardware (e.g., Youboxun i9000 [7]). Based on the localization module, multi-modal sensing data is generated when couriers are working carrying PDAs. The data is summarized in Table 1.

Table 1. Multi-modal Sensing Data

	Trajectory	Sensor Data	Reporting
Feature	Latitude, Longitude, Timestamp	Light intensity, Cellular signal, Magnetic field	OrderID, Timestamp
Sampling	10 seconds for 1 point	1 second for 1 point	Uncertain

Trajectory. For communities in metropolitan areas with high buildings and trees, there are significant noises brought by environmental effects and device hardware issues that will impact localization accuracy. To improve the localization accuracy and quality of trajectories, we design a localization strategy based on WiFi and GPS. We use the raw GPS measurement for localization by default. Specifically, we estimate the quality of the GPS signal by the *Accuracy* parameter. When the *Accuracy* is greater than a threshold (e.g., 100 meters), we switch to utilize the network localization approach provided by commercial mapping services [44, 60, 78]. As shown in Table 1, the GPS is sampled every 10 seconds and a group of consecutive GPS points form a trajectory $Traj = \langle p_1, p_2, \dots, p_i, \dots, p_n \rangle$, where $p_i = (lat_i, lng_i, t_i)$, representing the latitude, longitude, and timestamp. These trajectories can be used to extract topology and geometry information of roads.

Sensor data. During the working process, the PDAs collect data from different sensors. As shown in Table 1, the sensor data include three different types of sensors in PDAs, i.e., light intensity, cellular signal, and magnetic field

sensors. These sensor data can reflect the couriers' current status, e.g., indoors or outdoors. The light sensor, magnetic field sensor, and cellular sensor are sampled every 1 second for indoor and outdoor detection.

Courier reporting sensing data. During the working process, couriers manually report parcel statuses on PDAs, i.e., human reporting, an action performed by the courier to report a parcel is delivered at the user's address, denoted as $Rep = (lat, lng, t, orderID)$, representing the parcel $orderID$ is delivered at timestamp t at location (lat, lng) . Intuitively, human reporting represents both parcel status and courier status (arrive at customers' locations indoors). Therefore, we might utilize it for community road network generation.

We also investigate the battery consumption of the localization module. Results show that a regular PDA device with an Android 7 system consumes 0.64% of battery in 30 minutes, which is acceptable in real-world usage. Utilizing this localization module, we could collect useful sensing data from the delivery process for community road network generation.

3.2 Opportunities in Couriers' Sensing Data

For business accounting purposes, logistics companies collect massive courier mobility data [28, 65]. In this subsection, we conduct a data-driven investigation to show the characteristics of the courier mobility data and demonstrate the feasibility of generating community road networks based on such data.

(i) Frequent visiting patterns. In last-mile delivery, couriers visit communities multiple times every working day, and one community is visited by one or multiple couriers in a day, as shown in Fig. 3(a). Such frequent visiting pattern generates massive mobility data, which are recorded by mobile devices (i.e., PDAs). This massive mobility data benefits the generation of road networks because each visit is a sensing of partial community roads and it records the road topology and connectivity. As shown in Fig. 3(b), we also found that couriers are in communities for parcel delivery or parcel pick-up during most of their working times.

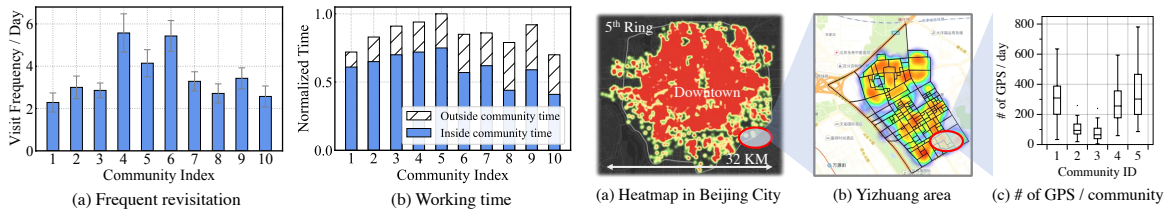


Fig. 3. Community visiting patterns.

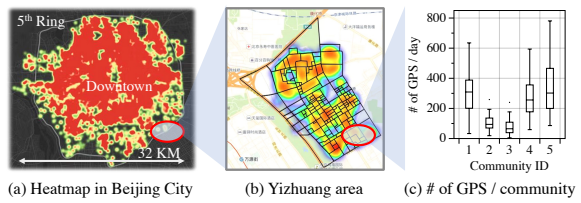


Fig. 4. Penetration rate in Beijing city.

(ii) High penetration rates. The platform we work with has more than 120,000 delivery couriers to serve over 500 million active users in over 600 cities. These couriers frequently travel in communities for parcel delivery or pickup, which brings a high community penetration rate. We sample 190,000 communities across all Chinese cities and find that the trajectory coverage rate is 96.3% (i.e., 96.3% of communities have been visited by couriers). Fig. 4(a) illustrates the coverage of couriers' mobility coverage in the Chinese city of Beijing on a normal workday, and we can find couriers' mobility covers most parts of the city, i.e., most communities. Downtown areas are visited more often due to the high community density with more parcels delivered. We further zoom in and present the heatmap of a region in Fig. 4(b). There are typically multiple GPS points every day in each community generated during couriers' parcel delivery process as shown in Fig. 4(c).

3.3 Challenges for Utilizing Couriers' Sensing Data

However, it is still nontrivial to generate community road networks with the above opportunities.

(i) **Uncertain courier behavior.** Unlike vehicle trajectories on open roads, delivery couriers' trajectories are more complex due to behavior uncertainty. Couriers' trajectories in communities consist of driving, walking, and staying. Also, couriers consecutively switch between indoor and outdoor statuses for the delivery as shown in Fig. 5. Trajectories collected during staying status may have huge GPS measurement errors due to the GPS hardware characteristics and environmental effects, e.g., the GPS points will still change when the courier is staying at a fixed location. The trajectories in an indoor context do not contribute to road generation as they are not on roads and might bring negative effects. Therefore, how to detect courier status and filter unnecessary trajectories is important but challenging.

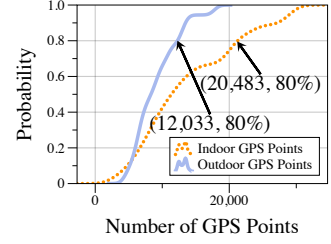


Fig. 5. Uncertain mobility.

(ii) **Inaccurate sensing.** There are usually more high buildings and trees in communities as shown in Fig. 1(c), which potentially brings high GPS localization errors. As shown in Fig. 6a, we found 50% of GPS points have sampling intervals of more than 10 seconds, which is caused by three main reasons: (i) weak GPS signals in community areas; (ii) the localization module is killed by the operating system; (iii) the device is shut down due to low battery. To measure the GPS deviation in communities, we carry two devices: a high-precision device equipped with an RTK (Real Time Kinematic) module and a standard PDA utilized by couriers. The ground truth is derived from the RTK device (deviation error is less than 1 meter) and manual inspection. We calculate the difference between GPS points from two devices at the same timestamp. Results in Fig. 6b show that for in-community settings, more than 50% points have deviation errors greater than 17.8 meters. In addition to environmental factors, GPS also deviates from the precise location when switching between GPS localization and network localization strategies.

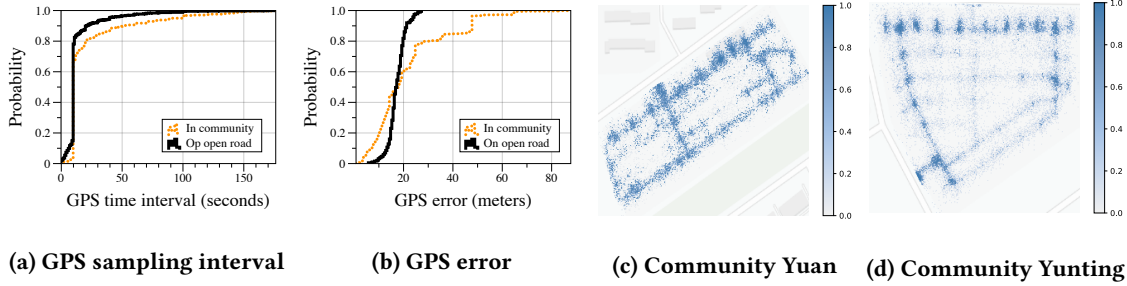


Fig. 6. Uncertainty in (a) sampling interval and (b) measurement error. (c)(d) Examples of unbalanced sensing on community roads.

(iii) **Unbalanced sensing.** Couriers' mobility in communities is unbalanced, i.e., some roads are visited with high frequency and some others are barely visited, which brings uneven GPS coverage rates on different road segments and makes it challenging to sense community road networks. Examples are shown in Fig. 6d and Fig. 6c.

To address these challenges, we design SmallMap, a novel community road network generation framework based on couriers' multi-modal sensing data. Details will be introduced in the following section.

4 FRAMEWORK OF SMALLMAP

In this section, we introduce the framework of SmallMap. Fig. 7 shows the two primary components of SmallMap, i.e., the Trajectory of Interest (ToI) Detection module and the Community Road Network Generation module.

(i) **Trajectory of Interest (ToI) Detection:** This module is designed to detect ToIs, i.e., *trajectory segments* ($\langle p_i, \dots, p_j, \dots, p_{j+p}, \dots, p_{j+q}, \dots \rangle$) that cover community road networks when a courier is either walking or driving. The goal is to filter out noisy trajectories for effective road network generation.

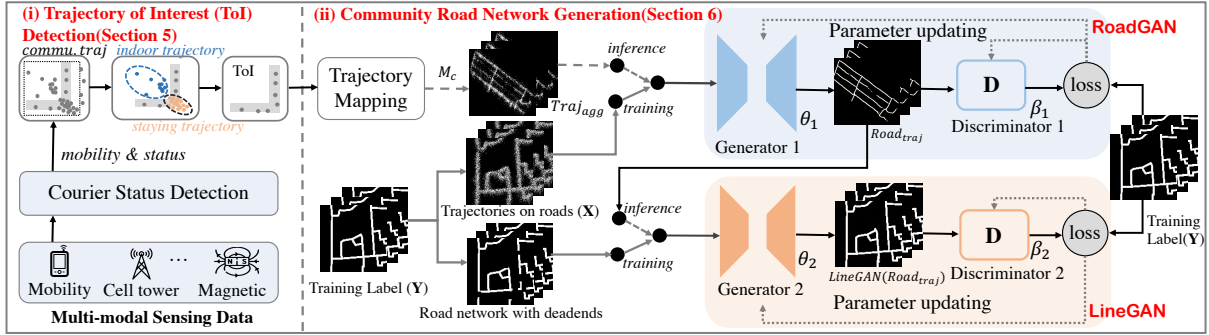


Fig. 7. Framework illustration. SmallMap consists of a ToI Detection module (in Sec. 5) and a Community Road Network Generation module (in Sec. 6).

(ii) **Community Road Network Generation:** Once trajectories on road networks have been identified, this module's objective is to determine the centerlines of noisy trajectories and generate interconnected roads through DualSTGAN, a hierarchical Generative Adversarial Network. DualSTGAN, trained with data generated by a novel data augmentation approach, is capable of generating community road networks from noisy in-community trajectories upon completion of the training phase.

5 TRAJECTORY OF INTEREST DETECTION

Courier trajectories comprise three components: walking or driving outdoors, remaining outdoors (e.g., resting on the grass), and staying or walking indoors. Moreover, these statuses are changing constantly. This is considerably different from the mobility patterns of business or private vehicles on urban roads. Even though the couriers' mobility does reflect the structure of community road networks, it is also mixed with various types of noise, i.e., indoor trajectories, trajectories that are not in communities, and couriers' stay patterns.

Such indoor or stay-pattern trajectories are not relevant to road network generation, and their removal can enhance the quality of the trajectory. Consequently, we design a two-step approach to detect and eliminate these trajectories, as illustrated in Algorithm 1. For each GPS point in the trajectory $Traj$, it will be added to the Trajectory of Interest (ToI) if both the indoor status detector f_{indoor} and staying status detector $f_{staying}$ indicate a negative result, and the in-community status detector $f_{community}$ indicates a positive result, i.e., the point is in communities and it is neither indoors nor in a static position.

5.1 Indoor and Outdoor Status Detection

Indoor trajectories fail to represent the road structure accurately, resulting in erroneous road network inferences. Therefore, we need to detect and remove these trajectories. Intuitively, we can easily segment indoor trajectories utilizing the spatial boundary of buildings, i.e., trajectories within building boundaries would be categorized as indoor trajectories. However, our empirical findings reveal that this approach is not feasible in a real-world setting, because indoor trajectories have large deviation errors and thereby often drift outside of building boundaries.

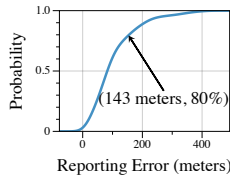
Another approach is reporting-based indoor trajectory detection. During delivery, couriers are required to report parcel delivery accurately when they deliver parcels at customers' locations inside of buildings. Thus, we can segment trajectories according to the reporting timestamps. However, we found that courier reporting is highly uncertain and unreliable as shown in Fig. 8(a), which shows the large deviation between reporting locations and the actual locations of couriers.

Algorithm 1: Trajectory of Interest Detection**Input:** Couriers' Trajectories $Traj$ **Output:** Trajectory of Interests ToI

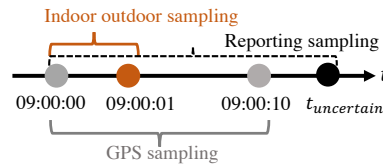
```

1  $ToI = []$ 
2  $i = 0$ 
3 while  $i < \text{len}(Traj)$  do
4    $status_1 \leftarrow f_{\text{indoor}}(Traj[i])$  # detect whether a point is in indoor
5    $status_2 \leftarrow f_{\text{community}}(Traj[i])$  # detect whether a point is in community
6    $status_3 \leftarrow f_{\text{staying}}(Traj[i])$  # detect whether a point is in staying
7   if not  $status_1$  and  $status_2$  and not  $status_3$  then
8      $ToI.add(Traj[i])$ 
9    $i \leftarrow i + 1$ 

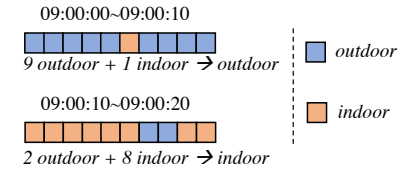
```



(a) Reporting distance error



(b) Sampling rate illustration



(c) Voting process illustration

Fig. 8. (a) Courier reporting errors, (b) Sampling rates of indoor outdoor status, GPS, and reporting, and (c) Voting process illustration.

The PDAs carried by couriers offer an opportunity to overcome the challenge of indoor/outdoor status detection by utilizing the sensors embedded in these devices. Accordingly, we design a status detection approach built upon IODetector [88, 89], a cutting-edge method for discerning between indoor and outdoor environments. In status detection, we harness multiple sensors (light intensity, cellular signal, and magnetic field measurements) to distinguish between outdoor and indoor statuses. The rationale is that indoor and outdoor environments differently impact the readings of these sensors. Hence, we can detect the couriers' status by modeling these sensor values. The final detection outcome is the average result from the three sensors.

One approach is to collect and relay sensor data to the server for detection. However, the data transmission brings substantial data and energy costs for PDAs, making it impractical for large-scale implementation. Therefore, we design a lightweight approach that can operate on PDAs and relay the detected results to the server, obviating the need to transmit raw sensor data. Nonetheless, our findings indicate that relying solely on sensor data is inadequate for our purposes due to environmental influences such as buildings and trees within community settings. To address this, we have developed a voting-based approach to improve the accuracy of our detection system. In particular, the detection algorithm is executed every second, which produces a series of 10 statuses for each trajectory point (given that the GPS sampling rate is set as one point per 10 seconds). The details of the sampling rate and the voting approach are depicted in Fig. 8(b) and Fig. 8(c), respectively. A trajectory point is deemed to be indoors if the majority of its status sequence suggests indoor statuses. Consequently, indoor trajectories can be detected, paving the way for subsequent road network generation.

5.2 In-community Status Detection

Following the detection of indoor trajectories, the subsequent step involves the detection of in-community trajectories. As shown in Fig. 9(a), a courier travels from an external location to a community, generating corresponding trajectories ($Traj_{raw}$). When trajectories are directly filtered using the Area of Interest (AoI) of community boundaries, drifting points p_1 , p_2 , and p_3 are mistakenly incorporated as in-community trajectory points. This inclusion adversely impacts the performance of road network generation. Therefore, we detect such drifting points to establish in-community trajectories ($Traj$) by computing each point's preceding and succeeding points. That is, drifting points are determined by their preceding and succeeding points being outside the community within a specific time window.

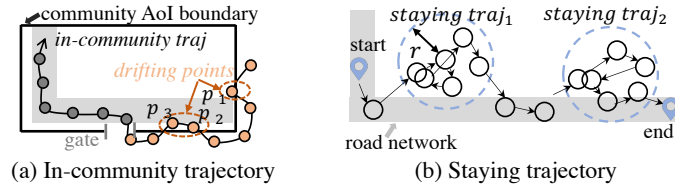


Fig. 9. In-community trajectory and staying status detection.

5.3 Staying Status Detection

In communities, couriers sometimes stay on the grass to take a break or contact customers (not on roads), as illustrated in Fig. 9(b). During these pauses, the GPS signals received can drift away from the true locations and cannot reflect the road networks, leading to considerable deviation errors if these trajectories are not detected. As shown in Fig. 10, the distance deviation in a stay point can be more than 100 meters. The number of GPS records and stay duration in a stay point are also large. For each trajectory segment in a day, we detect stay points (i.e., *groups of consecutive points in a trajectory that are produced during a stationary period of a moving object at a specific location*) based on the location and timestamp information [52]. Specifically, if a trajectory segment stays within *radius* meters for more than *m* minutes, the segment is identified as a stay point. The parameters *radius* and *m* can be adjusted according to different scenarios, such as setting a smaller *r* in areas with minimal GPS drifting. Based on our observations from real-world data, we set *radius* as 20 and *m* as 2 in our experiment.

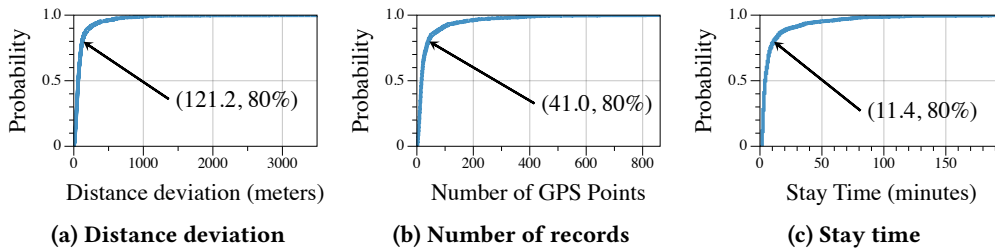


Fig. 10. Characteristics for each stay point.

After the detection of indoor and stationary statuses, we categorize the original courier trajectories into two groups: trajectories that cover community road networks (ToI), and trajectories that are not pertinent to road network generation. The computation of ToI is demonstrated in Eqn. 1.

$$ToI = Traj \setminus (Traj_{indoor} \cup Traj_{staying}) \quad (1)$$

where $Traj_{indoor}$ and $Traj_{staying}$ are the indoor trajectory set and stationary trajectory set, respectively. The set union operator $\cup$ combines the two subsets of trajectories that we aim to exclude from the $Traj$ into one set. The set difference operator $\setminus$ is used to subtract both $Traj_{indoor}$ and $Traj_{staying}$.

6 COMMUNITY ROAD NETWORK GENERATION

Motivation. The subsequent step involves generating road networks from the noisy and accumulated sparse trajectories. According to our data analysis, one critical observation is that the GPS signals follow the mixed Gaussian distributions, and the centerline of roads can be discerned by humans when the trajectories are displayed graphically. Thus, rather than generating road networks through various stages of trajectory processing, as in most trajectory-based road generation studies [27, 43, 54], we reframe this task as a translation problem. Recently, Generative Adversarial Networks (GANs) have been empirically evaluated in the sensing community and exhibit great performance in reconstructing accurate topology information from sensing signals [55]. Therefore, we design a GAN-based model to generate community road networks from distributively collected trajectories.

The right part of Fig. 7 presents our generative model, DualSTGAN, which is composed of two modules: RoadGAN and LineGAN, respectively. To tackle the challenge of training data shortage, we design a novel approach to augment road network data for model training. After the training phase, DualSTGAN generates community road networks from Trajectory of Interests as shown in Fig. 11.

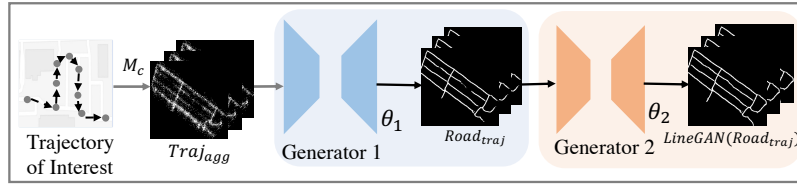


Fig. 11. Inference phase of DualSTGAN.

6.1 Road Network Topology Generation

The Generative Adversarial Network (GAN) is a representative generative model that has demonstrated excellent performance in various domains, including image generation and natural language generation [30]. A GAN consists of a generator (G) that synthesizes a sample from an input dataset (which can either be noise or a real sample from other domains), and a discriminator (D) that inputs data from either the generator or the real sample database and outputs a probability representing the likelihood of the input coming from the real sample database. The generator and discriminator are co-trained to achieve a Nash equilibrium [30], as shown in Eqn. 2.

$$\min_G \max_D L(D, G) = E_{y \sim P_r(y)} [\log(D(y))] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (2)$$

where y are sampled from the real dataset and z are sampled from random noises.

6.1.1 Domain Alignment. Nevertheless, the application of GAN for generating community road networks from spatiotemporal trajectories is not straightforward. Trajectories contain timestamp information, which differs from road network data. This domain gap between the two datasets poses a challenge to conventional GAN models, making it impractical to directly use trajectories as model input. Consequently, we design a mapping mechanism to project raw trajectories and road networks into the same domain for model training and inference.

Given a collection of the ToIs ($\langle ToI_1, ToI_2, \dots, ToI_n \rangle$) in a community, we project all trajectory points into a two-dimensional matrix $Traj_{agg}$ by normalizing trajectories using the spatial boundaries of the community, as shown in Eqn. 3.

$$Traj_{agg} = M_c(\langle ToI_1, ToI_2, \dots, ToI_n \rangle) \quad (3)$$

Here, $\mathcal{M}_c$ represents the mapping process and c is the information about community boundaries. Specifically, each GPS point in ToIs is mapped by calculating $\frac{lat-lat_{min}}{lat_{max}-lat_{min}} \times d_2$ and $\frac{lng-lng_{min}}{lng_{max}-lng_{min}} \times d_1$, where lat_{max} , lat_{min} , lng_{max} , and lng_{min} are the minimum and maximum latitude and longitude of the community. lat and lng are the latitude and longitude of the current GPS point. d_1 and d_2 denote the width and height of the matrix, and both of them are set to 256 in the experiment to match the model input size. $Tra_{agg} \in \mathbb{R}^{d_1 \times d_2}$, n is the number of *ToI* in the community. Consequently, all trajectory points are embedded in this two-dimensional matrix as the input of our GAN model to generate community road networks.

6.1.2 Generation Networks. Unlike generating a sample from noise, *conditional GAN (cGAN)* typically generates samples from source domain data samples. In road network generation, the source domain is noisy trajectories and the target domain is the road networks. In the conditional generation setting, the model can be guided more easily to generate road networks that follow the true geometry and topology as the training label. The loss function of conditional generation is in Eqn. 4.

$$L_{cGAN}(D, G) = E_{x,y}[\log(D(x, y))] + E_{x,z}[\log(1 - D(x, G(x, z)))] \quad (4)$$

where x are samples in the source domain, y are corresponding training labels in the target domain, and z are random noises.

The generative model in our setting is designed to map noisy trajectories into structured road networks. Therefore, we design a generator based on the encoder-decoder architecture. Specifically, the generator consists of two encoding blocks and two decoding blocks, with 6 residual blocks [35]. Each residual block contains a convolutional layer, a normalization layer, and an activation layer. As shown in Fig. 12, the encoder downsamples the input road network from 256×256 to 64×64 , and the residual blocks transform the features into the target domain. The input size is chosen because the majority of community widths in metropolitan areas are smaller than 500 meters, and 256 is a popular size following the experience from existing studies [63]. The decoder then upsamples the features and generates road networks of size 256×256 . To guide the generator in learning a correct mapping from trajectories to road networks, a CNN-based discriminator is utilized. The discriminator determines whether a patch of the generated road network is real or generated. Instead of predicting a whole road network to be real or generated, such patch-based discrimination provides fine-grained training supervision and it has shown superior performance in different tasks [46].

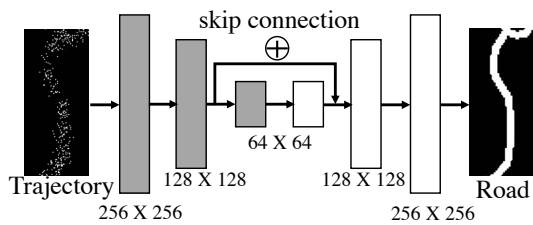


Fig. 12. Generator's encoder-decoder architecture. The input is noisy trajectories and the output is corresponding road networks.

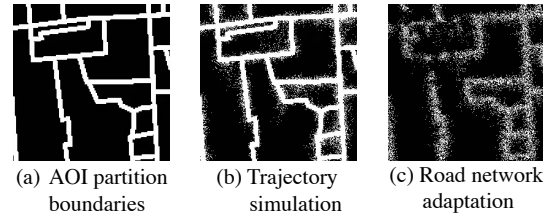


Fig. 13. Training data generation. (a) AOI partition boundaries as labels (Y). (b) Trajectory simulation. (c) Road network adaptation (X).

6.1.3 Training Data Construction. A key challenge in generating road networks with GANs is the shortage of training data, i.e., ground truth roads and noisy trajectories on corresponding roads (X). Intuitively, a road network partitions a region into different areas by road path, which is similar to Area-of-Interest (AoI) boundaries that partition an area into different AoIs. AoI boundaries consist of major street roads, building boundaries, and

other geographical elements such as rivers. Therefore, we might be able to train the road network generation model by utilizing existing AoI boundary data that is easy to obtain.

(i) AoI partition boundary generation (Y). AoI boundary data can be collected on platforms such as OpenStreetMap (OSM) [5], which are consecutive spatial points (*latitude, longitude*) representing the spatial boundary of a region. However, the main road network is very different from community roads, e.g., the shape is very different. Therefore, we combine AoI boundaries and main road skeletons by expanding AoI boundaries until they connect with a road to obtain synthetic roads, which are similar to community roads. In this way, we obtain synthetic road networks in various shapes as in Fig. 13 (a). These AoI boundaries can be utilized as the training label for road network generation.

(ii) Trajectory simulation. Even with synthetic community road networks as in Fig. 13 (a), we are still not ready to train our generation model. For real trajectories in communities, GPS points are spread alongside road segments that follow mixed Gaussian distributions. Therefore, we add noise to the AoI partition image (Fig. 13 (a)) by sampling synthetic GPS points from multiple Gaussian distributions as shown in Eqn. 5.

$$\hat{Traj}(r) = \sum_{c=1}^k \pi_c \mathcal{N}(r|\mu_c, \Sigma_c) \quad (5)$$

where $\hat{Traj}(r)$ represents simulated trajectories, k is the number of distributions, and $\sum_{c=1}^k \pi_c = 1$. The result is shown in Fig. 13 (b), which simulates the trajectories generated by PDAs.

(iii) Road network adaptation (X). The road network structure in Fig. 13 (b) is too clear and easy to recognize, which does not benefit the generative model as it cannot provide a realistic training environment that simulates the real trajectories. Thus, we conduct a two-step road network adaptation. Firstly, multiple noises are sampled centering each point on the road networks, and then we replace the original roads with these sampled points. This process blurs the road network skeleton, which is similar to real-world trajectories on community roads. Then, we observe that real-world trajectory points are not uniformly distributed on each road segment due to visiting frequency and preference variance. Based on this observation, we design a novel adaptation approach motivated by the pre-training task in the natural language processing community, where a few words in a sentence are masked and the model is trained to infer the masked content [53]. Specifically, we randomly mask a segment along the road networks and remove some noisy points to sparse the selected road segments, which generates road networks covered with simulated trajectories as in Fig. 13 (c). It can be observed that the generated trajectories are similar to the real trajectories on community road networks.

6.1.4 Model Training. Once we have constructed the training data (input **X** as in Fig. 13 (c) and training label **Y** as in Fig. 13 (a)), we can train RoadGAN to generate road networks from noisy trajectories $Traj_{agg}$.

$$Road_{traj} = RoadGAN_{\theta_1, \beta_1}(Traj_{agg}) \quad (6)$$

where $Road_{traj} \in \mathbb{R}^{d_1 \times d_2}$, θ_1, β_1 are trainable parameters for the generator and discriminator of RoadGAN, respectively. The training loss consists of two parts, i.e., loss from the conditional GAN, and the loss from L1 loss. Existing studies show that L1 loss captures low-level features such as boundary and shape information, which is important for road line generation. The loss is shown in Eqn. 7.

$$L_{RoadGAN} = L_{cGAN}(D, G) + \lambda L_{L1}(G) \quad (7)$$

where $L_{L1}(G) = E_{x,y,z}[\|y - G(x, z)\|_1]$ and λ is a hyperparameter that balances the importance of two losses.

6.2 Road Network Deadend Imputation

Motivation. One of the principal challenges faced by generative models in generating images with robust topological structures is the *deadend effect*. This refers to the abrupt discontinuation of generated lines or curves,

which adversely affects the performance of road network generation. This issue likely stems from the training mechanisms of generative models, wherein only portions of images are sampled throughout the training and optimization process [34]. Furthermore, the task of creating a coherent road network skeleton and seamlessly connecting road networks from noisy trajectories poses significant challenges for generative models due to the complexities inherent in the training process.

This phenomenon motivates us to design LineGAN to address the deadend issue and generate road networks with better topology and geometry. LineGAN is a generative model for road network imputation, i.e., infer incomplete or damaged segments of a road network based on partial roads. The generator and discriminator follow the same architecture of RoadGAN, but the training purpose is different, i.e., rather than training to generate road networks from noisy trajectories, LineGAN aims to impute broken deadends. However, it is challenging to train LineGAN due to the lack of training data. Thus, we design a novel data augmentation approach.

For each AoI boundary (i.e., training label (Y) as in Fig. 7), all roads are connected and there are no deadends. To simulate the results with deadends generated by RoadGAN, we first sample multiple grids in AoI boundaries and remove roads in the selected grids. Thus, we turn a connected AoI boundary into a road network that is similar to the road network generated by the RoadGAN, i.e., road networks with deadends as in Fig. 7. After training with the constructed data, LineGAN generates road networks $\hat{Road}$ based on the output of RoadGAN.

$$\hat{Road} = \text{LineGAN}_{\theta_2, \beta_2}(\text{Road}_{traj}) \quad (8)$$

where $\hat{Road} \in \mathbb{R}^{d1 \times d2}$, θ_2, β_2 are trainable parameters for the generator and discriminator of LineGAN, respectively.

7 EVALUATION

In this section, we introduce the evaluation setup, baseline comparison, and system efficiency analysis.

7.1 Evaluation Setup

7.1.1 Implementation details. Our framework is implemented on a machine with 40GB memory, P40 GPU with 24GB memory, Python 3.6, and Pytorch 1.6. Our RoadGAN is adapted from Pix2Pix Model [46]. We set the batch size as 1, and the maximum training epoch as 150. The initial learning rate is set as 0.0002. All deep learning-based baseline models are trained at least 100 epochs to ensure model convergence.

7.1.2 Dataset. We collect a multi-modal dataset from 119 communities during the delivery process of 205 couriers at the logistics platform we work with from 2022.07.01 to 2022.08.31, including 12,336 trajectories and 3.16 million GPS points. The communities are in the Chinese cities of Beijing (urban areas, 112 communities, 196 couriers) and Xuancheng (suburban areas, 7 communities, 9 couriers). The community AoI polygons and general AoI partition polygons are provided by the company's spatial database, generated by standard geospatial data production process, e.g., combining satellite images, geographical structures such as open road lines and river boundaries, and manual inspection. We utilize the AoI partition polygons to construct 4,000 trajectory-road network pairs to train the road network generation models. We collect the community road network ground truth with the professional team at the logistics company we work with to evaluate our model performance. Since the scale of the Beijing Dataset is much larger, results are shown based on the Beijing Dataset unless noted otherwise.

7.1.3 Ground Truth. The community road network ground truth consists of two parts. For communities with clear satellite images, we manually label the ground truth with professional map annotation software. For the rest of the communities, we manually visit communities with high-accuracy localization devices equipped with an RTK localization module. After annotating key points in communities, we further post-processing the annotated data on map annotation software. For courier status ground truth collection, we develop an APP that runs our

indoor status recognition approach. The App also supports marking the timestamp of getting in and out of a building.

7.1.4 Metrics. We use both qualitative methods and quantitative metrics (i.e., precision, recall, F1, and road network difference (rd)) to evaluate the generated road networks. We follow the commonly-used evaluation approach [11] to evaluate both topology (measuring the interconnections) and geometries (measuring the geographical location) of generated road networks.

Qualitative evaluation. To have a straightforward comparison of different methods, we compare the performance of representative methods qualitatively by visualizing generated road networks from two communities. More similarities between the generated road networks and the ground truth represent better performance.

Road network difference (rd). To measure the detailed difference between generated road networks with ground truth road networks, we calculate the road difference rate (rd). rd measures the ratio of pixels in generated road networks that are different from ground truth road networks as in Eqn. 9, where $\hat{r}$ and r are generated road network image and ground truth road network image, respectively. $v(r_{i,j})$ is the pixel value at location (i, j) in r . $H(x)$ is a step function that transforms x into 0 or 1. $H(x)=0$ if $x=0$, and $H(x)=1$ if $x \neq 0$. The smaller rd indicates better performance.

$$rd = \frac{\sum_{1 \leq i, j \leq N} H(v(\hat{r}_{i,j}) - v(r_{i,j}))}{N^2} \quad (9)$$

Precision, recall, and F1. We follow the de-facto road network evaluation approach to calculate the precision, recall, and F1 [11], which refers to the proportion of the generated roads that accurately represent true roads, the proportion of roads that are correctly captured by the model, and the harmonic mean of precision and recall, respectively. Specifically, we randomly select n starting locations in ground truth roads and put a hole with a radius of r every distance d until a maximum distance threshold s is reached. Then, for generated roads, we follow the same n , d , and s to put a *marble* every distance d . A hole is successfully matched with a marble if the marble is within the r -radius of the hole. We can calculate the correctly matched holes to evaluate the precision, recall, and F1. $precision = \frac{\# \text{ of matched marbles}}{\# \text{ of marbles}}$, $recall = \frac{\# \text{ of matched holes}}{\# \text{ of holes}}$, and $F1 = \frac{2 \cdot precision \cdot recall}{precision + recall}$.

7.1.5 Baselines.

- TrajCluster [76] is a representative method to generate city road networks and has achieved good performance. It generates road networks by clustering GPS trajectories generated by vehicles traveling on road networks. It can also infer road attributes by mining trajectory properties.
- coMap [27] utilizes complementary vehicle mobility data to infer city-scale road network structure and road categories. The intuition is that an increased volume of mobility data offers more detailed information to underline road networks. We adapt coMap in our setting using mobility data from multiple couriers.
- Pix2Pix [46] is a generative model trained on paired data. We train pix2pix with road networks and apply the trained model to generate road networks from trajectories to compare with our framework. Pix2pix utilizes U-Net [62] as the generator and PatchGAN as the discriminator.
- CycleGAN [90] is a Generative Adversarial Network trained on non-paired data. It has been utilized to generate road networks from aerial images. We train CycleGAN with unpaired road network data and generate road networks from noisy trajectories. The generator and discriminator of CycleGAN are the same as in Pix2Pix.
- RoadGAN is a variant of our DualSTGAN by removing the deadend imputation module.

7.2 Overall Evaluation

We evaluate the model performance based on both qualitative evaluation and quantitative evaluation.

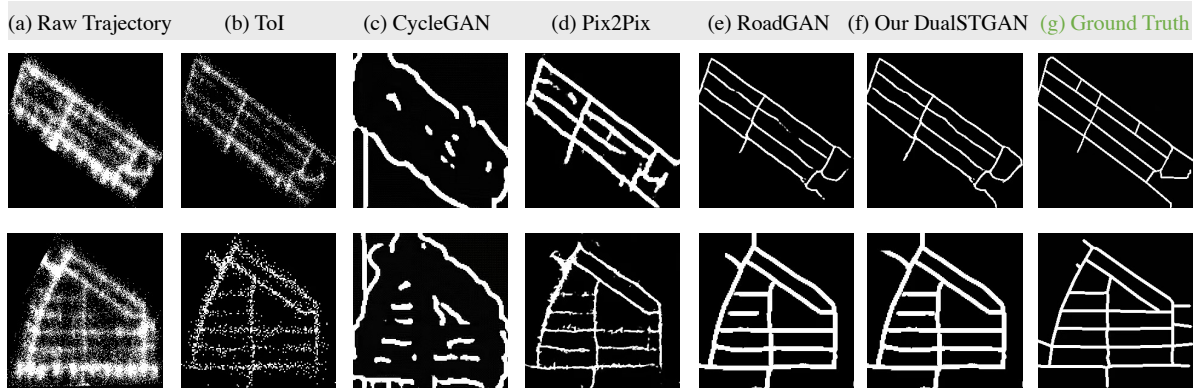


Fig. 14. Road network generation results. (a) is the raw data (couriers' GPS points in the communities). (b) shows the Trajectory of Interest (ToI) after removing staying and indoor trajectories. (c)-(f) shows the generation results by different methods. (g) is the road network ground truth.

Qualitative evaluation. We compare different models by the visual differences between generated roads with the ground truth roads in Fig. 14. It can be noted that our SmallMap generates roads most similar to the ground truth roads in terms of overall topology and geometry. Moreover, with LineGAN designed to infer deadends, the interconnection of generated roads is improved significantly. Another interesting observation is that even trained with the same generator and discriminator architecture, CycleGAN generates much worse road networks than Pix2Pix. This phenomenon demonstrates that unsupervised training is hard to yield very good performance in translating noisy trajectories into corresponding road networks. It is probably because unsupervised training in CycleGAN guides the model to learn the road network "style", i.e., road networks are connected lines rather than noisy points, but hard to provide enough training guidance to generate lines corresponding to noisy points and maintain geometry and topology information.

We also show the imputation effect of LineGAN in Fig. 15. LineGAN is trained to impute road segments in favor of road connectivity. It can be noted that the road network with multiple deadends (Fig. 15(a)) can be imputed as Fig. 15(b), which is very close to the ground truth road network (Fig. 15(c)).



Fig. 15. LineGAN imputation result.

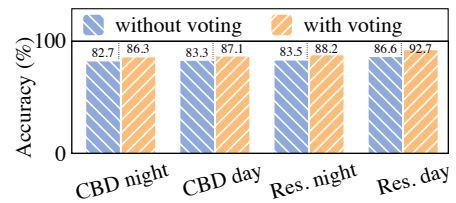


Fig. 16. Indoor context evaluation.

Quantitative evaluation. The comparison between SmallMap and baseline models is illustrated in Table 2. SmallMap achieves the best performance among all metrics in two cities with an average F1-score of 88.9%. Note that the recall in the Xuancheng Dataset is lower, which is probably because there are fewer orders in communities in Xuancheng and thereby fewer courier trajectories. According to our communication with the business team of the logistics platform, such performance can meet most of the business needs such as travel time estimation.

Overall, the results in the two cities are consistent. CycleGAN has the highest rd, which means road networks generated by CycleGAN are the most different from ground truth road networks, which follows the phenomenon in Fig. 14. One of the main reasons is the unpaired training paradigm in CycleGAN, i.e., the training data is not one-to-one paired data as in Pix2Pix. Therefore, it is challenging for CycleGAN to precisely translate trajectories to corresponding road networks. Pix2Pix, coMap, and TrajCluster have similar scores of rd but significantly different precision, recall, and F1 scores. This is probably because road networks generated by these methods are similar in terms of absolute difference from ground truth road networks, but are different in terms of road geometry and topology, which are captured by the F1-score. For trajectory-based models, TrajCluster and coMap, the inaccurate individual GPS trajectories might be the main reasons that decrease their performance. From RoadGAN to DualSTGAN, the performance is improved by imputing the broken deadends, which follows the visual comparison in Fig.14.

Table 2. Overall Evaluation Result

Model	Beijing Dataset				Xuancheng Dataset			
	rd	Precision	Recall	F1	rd	Precision	Recall	F1
CycleGAN	0.728	0.561	0.533	0.547	0.593	0.612	0.596	0.604
TrajCluster	0.155	0.631	0.685	0.657	0.168	0.626	0.674	0.649
coMap	0.142	0.735	0.761	0.748	0.149	0.718	0.701	0.709
Pix2Pix	0.136	0.832	0.810	0.821	0.141	0.813	0.797	0.805
Our RoadGAN	0.059	0.883	0.876	0.879	0.077	0.871	0.830	0.850
Our DualSTGAN	0.051	0.913	0.907	0.910	0.064	0.896	0.842	0.868

Indoor context detection evaluation. To evaluate the performance of indoor context detection, we implement our algorithm on PDA and Android smartphones. To collect the ground truth, we develop an APP, to record the time duration of indoor contexts. We compare the detection accuracy in different regions and time slots, shown in Fig. 16. The overall accuracy is above 80%. The accuracy of residential areas is higher because there are more complex indoor environments in Central Business District (CBD) areas. At the same time, the accuracy in the daytime is better than at nighttime. It can be noted that the accuracy is improved with the voting-based approach.

7.3 Impact of Different Components

To further investigate the importance of different components, we conduct experiments to evaluate their impacts on the performance of SmallMap.

Impact of deadend imputation. In this work, we develop LineGAN for the deadend imputation to improve the performance of community road network generation. The qualitative and quantitative results are shown in Fig. 14 and in Table 2, respectively. We evaluate how the deadend imputation component contributes to the quality of generated roads. Fig. 14 shows that with deadend imputation, the model can successfully identify the broken generated road segments and impute them with road lines. Quantitatively, the deadend imputation module improves the F1-score by 3.1% in Beijing City and 1.8% in Xuancheng City as shown in Table 2.

Impact of radius r . The hyper-parameter radius r determines the error tolerance level and thereby impacts the evaluation results. Intuitively, a larger r represents a more coarse evaluation granularity and brings higher precision and recall. The impacts of r on precision, recall, and F1-score are shown in Fig. 17, Fig. 18, and Fig. 19, respectively. Generally, the results of all models are improved with the increase of radius r , which follows our

assumption. Also, the performance gap among the five models decreases with the increase of r . This is because a large r matches a point on generated road networks (*marble*) with another point (*hole*) on ground truth road networks even though they are far away from each other.

Impact of the number of starting points. Intuitively, the location of starting point determines the locations of *marbles* and *holes*, which then impact precision, recall, and F1. The number of starting points impacts the variance of calculated metric values. Therefore, we investigate how variance changes along with the number of starting points by calculating the variance as shown in Fig. 20. As a demonstration, we compare the variance of CycleGAN and DualSTGAN, which represent low-performance and high-performance models, respectively. It can be noted that the variance of both models continuously decreases with the increase of the starting points, which is probably because more starting points cover more possible comparison points on generated and ground truth road networks, and thereby the result has a smaller variance. Also, the variance of CycleGAN is significantly higher than DualSTGAN before 30 starting points, because the connectivity of the generated road network by CycleGAN is worse. The variances become stable after 30, we thereby choose 30 starting points in our experiments.

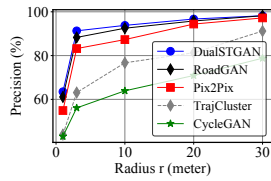


Fig. 17. Impact of radius on precision.

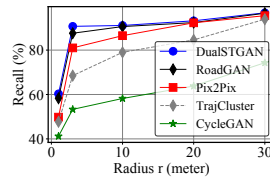


Fig. 18. Impact of radius on recall.

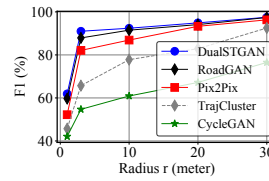


Fig. 19. Impact of radius on F1-score.

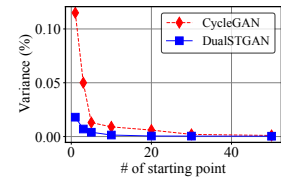


Fig. 20. Impact on variance.

Impact of road network adaptation. To verify the effectiveness of our training data augmentation process, we remove the road network adaptation in the trajectory-road training pair generation process. The results are shown in Table 3. Experiments show that without the adaptation step, the performance decreases significantly, which is probably because of the great difference between simulated trajectories and real-world couriers' trajectories. This analysis highlights the great importance of the third step of the training data generation process.

Table 3. Impact of Road Network Adaptation Module

Model	rd	Precision	Recall	F1
w/o adaptation	0.153	0.742	0.720	0.731
Our DualSTGAN	0.051	0.913	0.907	0.910

7.4 Post-processing of Road Networks

Based on the generated road network structures, we can further generate road graphs by detecting road segments in the road networks [22]. We show examples in Fig. 21. Note that, road segments can be further processed and merged to reduce the storage cost, which is not the focus of this paper.

7.5 Computation Efficiency Analysis

Efficiency is an important factor that impacts the large-scale application capability of SmallMap. Therefore, we analyze the efficiency of SmallMap in terms of training speed and inference speed. The training of DualSTGAN takes 2.8 hours on a single P40 GPU for 100 epochs with 4,000 training images. For a city of size 50km $\times$ 50km, we can partition it into 500 meters $\times$ 500 meters grids, which yield 10,000 grids in total. Then, we can generate the road networks for the whole city in 1.9 hours, with 0.7 seconds for each grid. Note that larger grid sizes bring

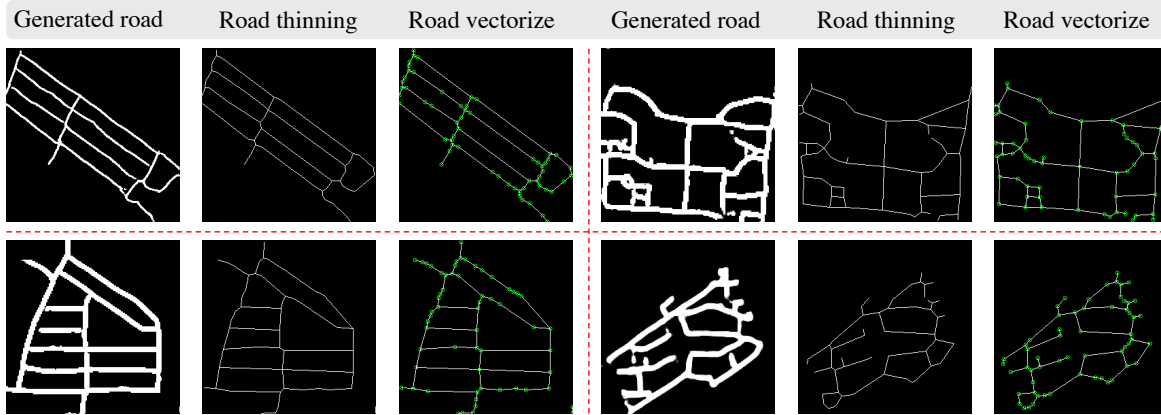


Fig. 21. Post-processing of generated community road networks.

fewer grids to be generated and thereby leading to a shorter generation time. This demonstrates the efficiency of our method for generating large-scale community road networks.

8 REAL-WORLD CASE STUDIES

We conduct real-world case studies in a region around the fifth ring of Beijing City, China, to support important applications in last-mile delivery utilizing our generated community road networks.

8.1 Courier Workload Estimation

In last-mile delivery, an important task is to estimate the workload for delivering a group of parcels, which benefits various tasks such as equitable assignment of tasks or rewards. Calorie consumption, which can be measured by commercial bracelets [4], can be utilized to quantify the difficulty level of delivery tasks. However, due to high costs and privacy concerns, the bracelet-based method cannot be applied on a large scale. Instead, a straightforward approach is estimating workload based on the number of delivered parcels, i.e., assuming more parcels bring a higher workload. However, due to the significant environmental differences in metropolitan areas, the difficulty of delivering parcels can vary significantly. For example, delivering 10 parcels to one address might be easier than delivering 5 parcels to 5 addresses that are far away from each other. Therefore, we design a new method as follows.

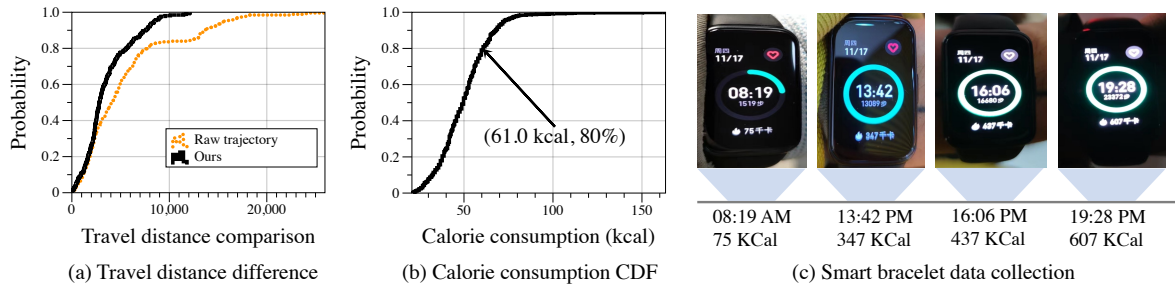


Fig. 22. Smart bracelet illustration. (a) shows a substantial discrepancy in couriers' travel distance between the result calculated from raw trajectories and our method. (b) shows couriers' calorie consumption patterns. (c) shows an example of our smart bracelet recording couriers' calorie consumption.

Method. The most relevant features related to couriers' workload include the horizontal moving distance, the weight of parcels, etc. Most horizontal moving distance is generated in communities, which rely on fine-grained community road networks for the measurement. Therefore, based on SmallMap, we design a technical framework to estimate the workload of couriers. Specifically, we design a learning-based method to estimate the calorie consumption of delivery couriers from a set of features. The features include *parcel weight*, *parcel volume*, *walk distance*, *work time span*, and *number of parcels*, etc. *walk distance* is calculated by map matching [57] couriers' trajectories with the generated road networks.

Ground truth collection. For model training and evaluation, we design a ground truth collection system based on smart bracelets (i.e., Keep B2 Bracelet [4]), which is illustrated in Fig. 22. It has been widely adopted to utilize wearable devices for calorie consumption measurement [48]. Human status detection based on smart bracelets has been widely studied and some applications such as sleep monitoring have been utilized in commercial devices [4, 49, 67]. We collaborate with 18 couriers in the logistics platform we work with and collect the calorie consumption data from 2023.02.17 to 2023.03.14. The hourly calorie consumption is shown in Fig. 22(b) by wearing smart bracelets as in Fig. 22(c). Each courier receives 30 RMB per day as a reward for wearing the bracelet during working time to collect the ground truth data.

Results. The result is shown in Table 4 using a linear regression (LG) model and gradient-boosted decision tree (GBDT) model [29]. For methods without generated road networks (w/o SmallMap), the estimated walking distance is significantly different from the actual distance as shown in Fig. 22(a), which potentially brings negative impacts on calorie estimation. The superior performance with SmallMap follows our assumptions on the importance of walking distance and our method.

Table 4. Calorie Estimation Results

	Method	MAPE	MAE	RMSE
w/o SmallMap	LG	0.23	10.34	13.93
	GBDT	0.22	10.09	14.03
with SmallMap	LG	0.19	8.91	12.28
	GBDT	0.17	7.97	10.47

8.2 Travel Time Estimation

Estimating the travel time between couriers' current locations with other locations is one of the most important tasks in delivery services, which enables many applications such as customer experience improvement and courier scheduling strategies [84]. The most important factors that impact travel time include travel distance and travel speed. During parcel delivery, couriers travel in communities mainly by walking and travel on open roads by driving delivery vehicles, both with relatively stable speeds. Therefore, the accurate travel distance between the two locations has a huge impact on the estimation of travel time.

We utilize the shortest path query method [47] on road network data to get the travel distance. According to the walking and driving speeds depicted in Figure 23a, it is evident that the majority of couriers' speeds are less than 5.8 km/hour and 18.6 km/hour, respectively. Thus, we set the walking speed as 5 km/hour and the driving speed as 18 km/hour, similar to other research studies [6]. In Fig. 23b, we present the error rate of travel time estimation based on the generated community road networks by SmallMap. It can be noted that, without community road networks (w/o *community road*), the error rate is the highest (mean error rate is 0.55, i.e., the ratio of estimated travel time error to ground truth travel time). Also, utilizing lines to represent the actual travel distance (*line-based*) brings sub-optimal results (mean error rate is 0.303). With SmallMap, the mean error rate (0.092) is the lowest. The results demonstrate the utility of SmallMap for travel time estimation.

8.3 Fine-grained Order Assignment

In last-mile delivery [64] and on-demand delivery [88], one of the fundamental tasks is order assignment, i.e., assigning orders to couriers according to spatial areas [32]. Most existing methods focus on the assignment at

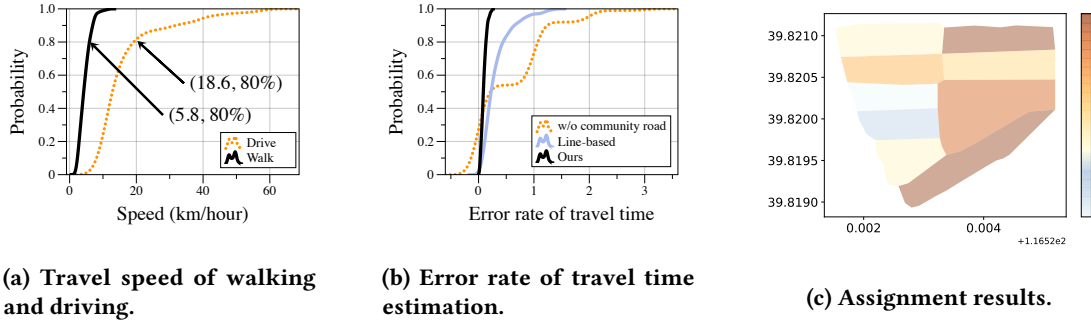


Fig. 23. (a) Travel speed of walking and driving, (b) Travel time error, and (c) Order assignment results.

the coarse-grained AoI level [32], which, while effective for most business operations, may not meet the specific business needs that require a more granular approach. To fill this gap, we propose a new method for fine-grained order assignment that leverages the generated community road networks from SmallMap.

Specifically, we first partition the community into fine-grained sub-areas according to the generated road networks from SmallMap. Then, we project each online order into one of the sub-areas by transforming its shipping address into a spatial location utilizing Geocoding services [28]. We utilize one-week (2023.12.10 to 2023.12.16) order data from the platform we work with and show the assignment results of a residential area in Fig. 23c, where darker colors indicate more orders in a sub-area, with the x-axis denoting longitude and the y-axis denoting latitude. Based on such fine-grained assignment results, the platform can conduct more fine-grained business, e.g., more efficient courier scheduling when multiple couriers work in the same community.

9 DISCUSSIONS

9.1 Lessons Learned

- **Low-cost crowdsensing for urban applications.** Urban sensing involves strong spatio-temporal characteristics and thereby brings the huge costs to maintain an up-to-date sensing system, e.g., a road network generation system. Designing a crowdsensing system with an incentive to involve citizens or workers would significantly lower the cost and also improve urban data quality for various urban applications such as last-mile delivery and instant food delivery. At the same time, these crowd-sensing data have different uncertainties, which need to be captured and addressed. Based on large-scale multi-modal sensing data, we have the opportunity to eliminate such uncertainties to improve the performance of urban applications such as community road network generation.
- **Collective human mobility.** Even though individual mobility has significant random noises due to environmental and contextual effects, collective mobility, however, has the potential to address the noises by fusing useful information from all individual mobility. Currently, human mobility data are increasingly available. It is important to use these collective human mobility data for ubiquitous computing towards low-cost sensing and high-quality urban services.

9.2 Limitations and Future Work

While SmallMap generates fine-grained community road networks with a low cost, we identify a few limitations that can be improved in future studies. Road network semantics (e.g., walking-only roads) are also important for real-world applications, which are not addressed in this study. Also, road networks generated by SmallMap still do not match the ground truth perfectly. The integration of more GPS trajectories and other data such as aerial images could potentially improve the road network sensing performance. Moreover, the evaluation scale is limited in the current study, and we plan to enhance the evaluation part in the future.

- **Road network semantics sensing.** Trajectory data and other sensor data from couriers' PDAs contain rich information and reflect the physical status of couriers, e.g., walking or driving. Therefore, based on the generated roads by SmallMap, advanced classification models can be designed to infer the road network semantics such as walking only or can be driven through, by classifying couriers' trajectories with models like Transformers [53].
- **Generative AI for road network sensing.** Given the great success of generative AI like GPTs [75] and diffusion models [66] in natural language processing and image understanding, we aim to design proper generative models for road network sensing. A key direction of future work involves the fusion of trajectory data and aerial imagery to harness the advantage of both trajectories (fine granularity and free-from-shadowing effects) and aerial images (wide coverage). We plan to investigate how to tokenize trajectories and aerial images for large foundation models [75] to get road networks in corresponding areas as the output.
- **Evaluation enhancement.** In the future study, we intend to increase the scale of evaluation by collecting ground truth data from more cities. Additionally, we plan to explore further case studies, for instance, developing a navigation App leveraging community road networks, thereby enriching the applications of our work.

9.3 Privacy Protection and Data Consent

We discuss privacy protection approaches in this study. (i) The data is collected from couriers' working devices, i.e., PDAs, which only record data during working hours. (ii) Couriers can turn off the localization function whenever they want. (iii) We only utilized aggregated mobility data in communities for road generation, and no personal information was accessed or used during the study.

10 CONCLUSION

In this paper, we design a novel framework called SmallMap to generate community road networks based on multi-modal sensing data. Specifically, SmallMap consists of two major modules, i.e., a Trajectory of Interest Detection Module and a Community Road Network Generation Module. To generate road networks with accurate topology structures, we design DualSTGAN, a hierarchical Generative model framework to generate road network skeletons and impute road network deadends. To tackle the challenge of training label shortage, we design a novel data augmentation approach to facilitate the training of DualSTGAN. Extensive results show that SmallMap outperforms state-of-the-art methods significantly. Moreover, we conduct concrete case studies to estimate the workload of delivery couriers, travel time, and order assignment by utilizing generated community road networks, which indicates the great potential of our SmallMap for real-world applications.

11 ACKNOWLEDGMENTS

This work is partially supported by the National Science Foundation under Grant No.2047822, 1952096, and 1951890. We thank all the reviewers for their insightful feedback to improve this paper.

REFERENCES

- [1] 2023. Baidu Maps. 2023. <https://maps.baidu.com/>.
- [2] 2023. Gaode Maps. 2023. <https://maps.gaode.com/>.
- [3] 2023. Google Maps. 2023. <https://maps.google.com/>.
- [4] 2023. Keep B2 Bracelet. 2023. <https://www.gotokeep.com/aiot>.
- [5] 2023. OpenStreetMap (OSM). 2023. <https://www.openstreetmap.org/>.
- [6] 2023. Scikit mobility. 2023. <https://scikit-mobility.github.io/scikit-mobility/>.
- [7] 2023. Youboxun i9000. 2023. <https://en.uovo.com/products/payment/I9000S.html>.
- [8] Ismi Abidi, Ishan Nangia, Paarijaat Aditya, and Rijurekha Sen. 2022. Privacy in Urban Sensing with Instrumented Fleets, Using Air Pollution Monitoring As A Usecase. In *Network and Distributed Systems Security Symposium*.

- [9] Muhammad Tayyab Asif, Nikola Mitrovic, Lalit Garg, Justin Dauwels, and Patrick Jaillet. 2013. Low-dimensional models for missing data imputation in road networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 3527–3531.
- [10] Favyen Bastani and Samuel Madden. 2021. Beyond road extraction: A dataset for map update using aerial images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 11905–11914.
- [11] James Biagioni and Jakob Eriksson. 2012. Inferring Road Maps from Global Positioning System Traces. *Transportation Research Record: Journal of the Transportation Research Board* 2291 (12 2012), 61–71. <https://doi.org/10.3141/2291-08>
- [12] Cheng Bo, Xiang-Yang Li, Taeho Jung, Xufei Mao, Yue Tao, and Lan Yao. 2013. SmartLoc: Push the Limit of the Inertial Sensor Based Metropolitan Localization Using Smartphone. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking* (Miami, Florida, USA) (*MobiCom '13*). Association for Computing Machinery, New York, NY, USA, 195–198. <https://doi.org/10.1145/2500423.2504574>
- [13] Rene Bruntrup, Stefan Edelkamp, Shahid Jabbar, and Bjorn Scholz. 2005. Incremental map generation with GPS traces. In *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005*. IEEE, 574–579.
- [14] Chu Cao, Zhidan Liu, Mo Li, Wenqiang Wang, and Zheng Qin. 2018. Walkway discovery from large scale crowdsensing. In *2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 13–24.
- [15] Lili Cao and John Krumm. 2009. From GPS traces to a routable road map. In *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*. 3–12.
- [16] Chen Chen, Cewu Lu, Qixing Huang, Qiang Yang, Dimitrios Gunopulos, and Leonidas Guibas. 2016. City-Scale Map Creation and Updating Using GPS Collections. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) (*KDD '16*). Association for Computing Machinery, New York, NY, USA, 1465–1474. <https://doi.org/10.1145/2939672.2939833>
- [17] Chen Chen, Cewu Lu, Qixing Huang, Qiang Yang, Dimitrios Gunopulos, and Leonidas Guibas. 2016. City-scale map creation and updating using GPS collections. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1465–1474.
- [18] Dongyao Chen and Kang G. Shin. 2019. TurnsMap: Enhancing Driving Safety at Intersections with Mobile Crowdsensing and Deep Learning. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 78 (sep 2019), 22 pages. <https://doi.org/10.1145/3351236>
- [19] Kongyang Chen and Guang Tan. 2019. BikeGPS: Localizing shared bikes in street canyons with low-level GPS cooperation. *ACM Transactions on Sensor Networks (TOSN)* 15, 4 (2019), 1–28.
- [20] Longbiao Chen, Chenhui Lu, Fangxu Yuan, Zhihan Jiang, Leye Wang, Daqing Zhang, Ruixiang Luo, Xiaoliang Fan, and Cheng Wang. 2021. UVLens: Urban Village Boundary Identification and Population Estimation Leveraging Open Government Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 57 (jun 2021), 26 pages. <https://doi.org/10.1145/3463495>
- [21] Xi Chen, Xiaopei Wu, Xiang-Yang Li, Yuan He, and Yunhao Liu. 2014. Privacy-preserving high-quality map generation with participatory sensing. In *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*. IEEE, 2310–2318.
- [22] Yao-Yi Chiang and Craig A Knoblock. 2010. Extracting road vector data from raster maps. In *Graphics Recognition. Achievements, Challenges, and Evolution: 8th International Workshop, GREC 2009, La Rochelle, France, July 22-23, 2009. Selected Papers 8*. Springer, 93–105.
- [23] Dragos Costea, Alina Marcu, Emil Slusanschi, and Marius Leordeanu. 2017. Creating roadmaps in aerial images with generative adversarial networks and smoothing-based optimization. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2100–2109.
- [24] JianXun Cui, Feng Liu, Davy Janssens, Shi An, Geert Wets, and Mario Cools. 2016. Detecting urban road network accessibility problems using taxi GPS data. *Journal of Transport Geography* 51 (2016), 147–157.
- [25] Anoooshmita Das, Emil Stubbe Kolvig-Raun, and Mikkel Baun Kjærgaard. 2020. Accurate Trajectory Prediction in a Smart Building Using Recurrent Neural Networks. In *UbiComp-ISWC'20*. ACM, 619–628.
- [26] Stefan Edelkamp and Stefan Schrödl. 2003. Route planning and map inference with global positioning traces. In *Computer science in perspective*. Springer, 128–151.
- [27] Zhihan Fang, Guang Wang, Xiaoyang Xie, Fan Zhang, and Desheng Zhang. 2021. Urban Map Inference by Pervasive Vehicular Sensing Systems with Complementary Mobility. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–24.
- [28] George Forman. 2021. Getting your package to the right place: supervised machine learning for geolocation. In *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part IV 21*. Springer, 403–419.
- [29] Jerome H Friedman. 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics* (2001), 1189–1232.
- [30] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (2020), 139–144.
- [31] Joachim Gudmundsson, Martin P Seybold, and John Pfeifer. 2021. On Practical Nearest Sub-Trajectory Queries under the Fréchet Distance. In *Proceedings of the 29th International Conference on Advances in Geographic Information Systems*. 596–605.

- [32] Baoshen Guo, Shuai Wang, Haotian Wang, Yunhuai Liu, Fanshuo Kong, Desheng Zhang, and Tian He. 2023. Towards Equitable Assignment: Data-Driven Delivery Zone Partition at Last-Mile Logistics. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*. Association for Computing Machinery, New York, NY, USA, 4078–4088. <https://doi.org/10.1145/3580305.3599915>
- [33] Baoshen Guo, Weijian Zuo, Shuai Wang, Wenjun Lyu, Zhiqing Hong, Yi Ding, Tian He, and Desheng Zhang. 2022. WePos: Weak-Supervised Indoor Positioning with Unlabeled WiFi for On-Demand Delivery. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 54 (jul 2022), 25 pages. <https://doi.org/10.1145/3534574>
- [34] Stefan Hartmann, Michael Weinmann, Raoul Wessel, and Reinhard Klein. 2017. Streetgan: Towards road network synthesis with generative adversarial networks. (2017).
- [35] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- [36] Songtao He, Favyen Bastani, Sofiane Abbar, Mohammad Alizadeh, Hari Balakrishnan, Sanjay Chawla, and Sam Madden. 2018. RoadRunner: improving the precision of road network inference from GPS trajectories. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 3–12.
- [37] Suining He and Kang G Shin. 2020. Towards fine-grained flow forecasting: A graph attention approach for bike sharing systems. In *WWW 2020*. 88–98.
- [38] Tianfu He, Jie Bao, Ruiyuan Li, Sijie Ruan, Yanhua Li, Chao Tian, and Yu Zheng. 2018. Detecting Vehicle Illegal Parking Events using Sharing Bikes' Trajectories.. In *KDD*. 340–349.
- [39] Zhiqing Hong. 2020. Generating location data with generative adversarial networks for sensing applications: PhD forum abstract. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems (Virtual Event, Japan) (SenSys '20)*. Association for Computing Machinery, New York, NY, USA, 800–801. <https://doi.org/10.1145/3384419.3430579>
- [40] Zhiqing Hong, Guang Wang, Wenjun Lyu, Baoshen Guo, Yi Ding, Haotian Wang, Shuai Wang, Yunhuai Liu, and Desheng Zhang. 2022. CoMiner: Nationwide Behavior-Driven Unsupervised Spatial Coordinate Mining from Uncertain Delivery Events. In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '22)*. Association for Computing Machinery, New York, NY, USA, Article 10, 10 pages. <https://doi.org/10.1145/3557915.3560944>
- [41] Zhiqing Hong, Haotian Wang, Wenjun Lyu, Hai Wang, Yunhuai Liu, Guang Wang, Tian He, and Desheng Zhang. 2023. Urban-scale POI Updating with Crowd Intelligence. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (<conf-loc>, <city>Birmingham</city>, <country>United Kingdom</country>, </conf-loc>)* (CIKM '23). Association for Computing Machinery, New York, NY, USA, 4631–4638. <https://doi.org/10.1145/3583780.3614724>
- [42] Zhiqing Hong, Heng Yang, Haotian Wang, Wenjun Lyu, Yu Yang, Guang Wang, Yunhuai Liu, Yang Wang, and Desheng Zhang. 2022. FastAddr: real-time abnormal address detection via contrastive augmentation for location-based services. In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems (<conf-loc>, <city>Seattle</city>, <state>Washington</state>, </conf-loc>)* (SIGSPATIAL '22). Association for Computing Machinery, New York, NY, USA, Article 64, 10 pages. <https://doi.org/10.1145/3557915.3560999>
- [43] Yourong Huang, Zhu Xiao, Xiaoyou Yu, Dong Wang, Vincent Havyarimana, and Jing Bai. 2019. Road network construction with complex intersections based on sparsely sampled private car trajectory data. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 13, 3 (2019), 1–28.
- [44] Mohamed Ibrahim, Ali Rostami, Bo Yu, Hansi Liu, Minitha Jawahar, Viet Nguyen, Marco Gruteser, Fan Bai, and Richard Howard. 2020. Wi-Go: Accurate and Scalable Vehicle Positioning Using WiFi Fine Timing Measurement. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services* (Toronto, Ontario, Canada) (MobiSys '20). Association for Computing Machinery, New York, NY, USA, 312–324. <https://doi.org/10.1145/3386901.3388944>
- [45] Sibren Isaacman, Richard Becker, Ramón Cáceres, Margaret Martonosi, James Rowland, Alexander Varshavsky, and Walter Willinger. 2012. Human mobility modeling at metropolitan scales. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. 239–252.
- [46] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1125–1134.
- [47] Donald B. Johnson. 1973. A Note on Dijkstra's Shortest Path Algorithm. *J. ACM* 20, 3 (jul 1973), 385–388. <https://doi.org/10.1145/321765.321768>
- [48] Ayanga Imesha Kumari Kalupahana, Ananta Narayanan Balaji, Xiaokui Xiao, and Li-Shiuan Peh. 2023. SeRaNDiP: Leveraging Inherent Sensor Random Noise for Differential Privacy Preservation in Wearable Community Sensing Applications. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 61 (jun 2023), 38 pages. <https://doi.org/10.1145/3596252>
- [49] Jiayu Li, Zhiyu He, Yumeng Cui, Chenyang Wang, Chong Chen, Chun Yu, Min Zhang, Yiqun Liu, and Shaoping Ma. 2022. Towards Ubiquitous Personalized Music Recommendation with Smart Bracelets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 125 (sep 2022), 34 pages. <https://doi.org/10.1145/3550333>

- [50] Jiawei Li, Linkun Lyu, Jia Shi, Jie Zhao, Junjie Xu, Jiuchong Gao, Renqing He, and Zhizhao Sun. 2022. Generating Community Road Network from GPS Trajectories via Style Transfer. In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems* (Seattle, Washington) (*SIGSPATIAL '22*). Association for Computing Machinery, New York, NY, USA, Article 24, 4 pages. <https://doi.org/10.1145/3557915.3560958>
- [51] Mingqian Li, Panrong Tong, Mo Li, Zhongming Jin, Jianqiang Huang, and Xian-Sheng Hua. 2021. Traffic Flow Prediction with Vehicle Trajectories. *AAAI* 35, 1 (May 2021), 294–302.
- [52] Quannan Li, Yu Zheng, Xing Xie, Yukun Chen, Wenyu Liu, and Wei ying Ma. 2008. Mining user similarity based on location history. In *In GIS '08: Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*. ACM, 1–10.
- [53] Yuxuan Liang, Kun Ouyang, Yiwei Wang, Xu Liu, Hongyang Chen, Junbo Zhang, Yu Zheng, and Roger Zimmermann. 2022. TrajFormer: Efficient Trajectory Classification with Transformers. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 1229–1237.
- [54] Xuemei Liu, James Biagioni, Jakob Eriksson, Yin Wang, George Forman, and Yanmin Zhu. 2012. Mining large-scale, sparse GPS traces for map inference: Comparison of approaches. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 669–677.
- [55] Chris Xiaoxuan Lu, Stefano Rosa, Peijun Zhao, Bing Wang, Changhao Chen, John A. Stankovic, Niki Trigoni, and Andrew Markham. 2020. See through Smoke: Robust Indoor Mapping with Low-Cost MmWave Radar. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services* (Toronto, Ontario, Canada) (*MobiSys '20*). Association for Computing Machinery, New York, NY, USA, 14–27. <https://doi.org/10.1145/3386901.3388945>
- [56] Yingtao Luo, Qiang Liu, and Zhaocheng Liu. 2021. STAN: Spatio-Temporal Attention Network for Next Location Recommendation. In *WWW 2021*. 2177–2185.
- [57] Wannes Meert and Mathias Verbeke. 2018. HMM with non-emitting states for Map Matching. In *European Conference on Data Analysis (ECDA)*, Date: 2018/07/04-2018/07/06, Location: Paderborn, Germany.
- [58] Mashaal Musleh, Sofiane Abbar, Rade Stanojevic, and Mohamed Mokbel. 2021. QARTA: an ML-based system for accurate map services. *Proceedings of the VLDB Endowment* 14, 11 (2021), 2273–2282.
- [59] Suraj Nair, Kiran Javkar, Jiahui Wu, and Vanessa Frias-Martinez. 2019. Understanding Cycling Trip Purpose and Route Choice Using GPS Traces and Open Data. *IMWUT*, 3, 1, Article 20 (March 2019), 26 pages.
- [60] Jiazhi Ni, Fusang Zhang, Jie Xiong, Qiang Huang, Zhaoxin Chang, Junqi Ma, BinBin Xie, Pengsen Wang, Guangyu Bian, Xin Li, and Chang Liu. 2022. Experience: Pushing Indoor Localization from Laboratory to the Wild. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking* (Sydney, NSW, Australia) (*MobiCom '22*). Association for Computing Machinery, New York, NY, USA, 147–157. <https://doi.org/10.1145/3495243.3560546>
- [61] Zhou Qin, Zhihan Fang, Yunhuai Liu, Chang Tan, Wei Chang, and Desheng Zhang. 2018. EXIMIUS: A Measurement Framework for Explicit and Implicit Urban Traffic Sensing. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (Shenzhen, China) (*SenSys '18*). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3274783.3274850>
- [62] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*. Springer, 234–241.
- [63] Sijie Ruan, Cheng Long, Jie Bao, Chunyang Li, Zisheng Yu, Ruiyuan Li, Yuxuan Liang, Tianfu He, and Yu Zheng. 2020. Learning to generate maps from trajectories. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 34. 890–897.
- [64] Sijie Ruan, Cheng Long, Zhipeng Ma, Jie Bao, Tianfu He, Ruiyuan Li, Yiheng Chen, Shengnan Wu, and Yu Zheng. 2022. Service Time Prediction for Delivery Tasks via Spatial Meta-Learning. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (Washington DC, USA) (*KDD '22*). Association for Computing Machinery, New York, NY, USA, 3829–3837. <https://doi.org/10.1145/3534678.3539027>
- [65] Sijie Ruan, Cheng Long, Xiaodu Yang, Tianfu He, Ruiyuan Li, Jie Bao, Yiheng Chen, Shengnan Wu, Jiangtao Cui, and Yu Zheng. 2022. Discovering Actual Delivery Locations from Mis-Annotated Couriers' Trajectories. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*. 3241–3253. <https://doi.org/10.1109/ICDE53745.2022.00307>
- [66] Chitwan Saharia, William Chan, Huiwen Chang, Chris Lee, Jonathan Ho, Tim Salimans, David Fleet, and Mohammad Norouzi. 2022. Palette: Image-to-Image Diffusion Models. In *ACM SIGGRAPH 2022 Conference Proceedings* (Vancouver, BC, Canada) (*SIGGRAPH '22*). Association for Computing Machinery, New York, NY, USA, Article 15, 10 pages. <https://doi.org/10.1145/3528233.3530757>
- [67] Florian Schaule, Jan Ole Johanssen, Bernd Bruegge, and Vivian Loftness. 2018. Employing Consumer Wearables to Detect Office Workers' Cognitive Load for Interruption Management. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 32 (mar 2018), 20 pages. <https://doi.org/10.1145/3191764>
- [68] Zhihao Shen, Wan Du, Xi Zhao, and Jianhua Zou. 2020. DMM: Fast map matching for cellular data. In *Proceedings of the 26th annual international conference on mobile computing and networking*. 1–14.
- [69] Rade Stanojevic, Sofiane Abbar, Saravanan Thirumuruganathan, Sanjay Chawla, Fethi Filali, and Ahid Aleimat. 2018. Robust road map inference through network alignment of trajectories. In *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM, 135–143.

- [70] Tao Sun, Zonglin Di, Pengyu Che, Chun Liu, and Yin Wang. 2019. Leveraging crowdsourced GPS data for road extraction from aerial imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7509–7518.
- [71] Tao Sun, Zonglin Di, and Yin Wang. 2018. Combining Satellite Imagery and GPS Data for Road Extraction. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery* (Seattle, WA, USA) (*GeoAI'18*). Association for Computing Machinery, New York, NY, USA, 29–32. <https://doi.org/10.1145/3281548.3281550>
- [72] Arvind Thiagarajan, Lenin Ravindranath, Hari Balakrishnan, Samuel Madden, and Lewis Girod. 2011. Accurate, Low-Energy Trajectory Mapping for Mobile Devices. In *Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation* (Boston, MA) (*NSDI'11*). USENIX Association, USA, 267–280.
- [73] Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts, Samuel Madden, Hari Balakrishnan, Sivan Toledo, and Jakob Eriksson. 2009. VTrack: Accurate, Energy-Aware Road Traffic Delay Estimation Using Mobile Phones. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems* (Berkeley, California) (*SenSys '09*). Association for Computing Machinery, New York, NY, USA, 85–98. <https://doi.org/10.1145/1644038.1644048>
- [74] Panrong Tong, Mingqian Li, Mo Li, Jianqiang Huang, and Xiansheng Hua. 2021. Large-scale vehicle trajectory reconstruction with camera sensing network. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*. 188–200.
- [75] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).
- [76] Yin Wang, Xuemei Liu, Hong Wei, George Forman, Chao Chen, and Yanmin Zhu. 2013. Crowdatlas: Self-updating maps for cloud and personal use. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. 27–40.
- [77] Jianzhong Yang, Xiaoqing Ye, Bin Wu, Yanlei Gu, Ziyu Wang, Deguo Xia, and Jizhou Huang. 2022. DuARE: Automatic Road Extraction with Aerial Images and Trajectory Data at Baidu Maps. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (Washington DC, USA) (*KDD '22*). Association for Computing Machinery, New York, NY, USA, 4321–4331. <https://doi.org/10.1145/3534678.3539029>
- [78] Yu Yang, Yi Ding, Dengpan Yuan, Guang Wang, Xiaoyang Xie, Yunhuai Liu, Tian He, and Desheng Zhang. 2020. TransLoc: Transparent Indoor Localization with Uncertain Human Participation for Instant Delivery. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking* (London, United Kingdom) (*MobiCom '20*). Association for Computing Machinery, New York, NY, USA, Article 41, 14 pages. <https://doi.org/10.1145/3372224.3419198>
- [79] Desheng Zhang, Jun Huang, Ye Li, Fan Zhang, Chengzhong Xu, and Tian He. 2014. Exploring human mobility with multi-source data at extremely large metropolitan scales. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. 201–212.
- [80] Hanyuan Zhang, Xingyu Zhang, Qize Jiang, Baihua Zheng, Zhenbang Sun, and Weiwei Sun. 2020. Trajectory similarity learning with auxiliary supervision and optimal matching. In *IJCAI'2020*. 11–17.
- [81] Ping Zhang, Zhifeng Bao, Yuchen Li, Guoliang Li, Yipeng Zhang, and Zhiyong Peng. 2018. Trajectory-driven influential billboard placement. In *KDD'2018*. 2748–2757.
- [82] Rui Zhang, Conrad Albrecht, Wei Zhang, Xiaodong Cui, Ulrich Finkler, David Kung, and Siyuan Lu. 2020. Map generation from large scale incomplete and inaccurate data labels. In *KDD'2020*. 2514–2522.
- [83] Rui Zhang, Conrad Albrecht, Wei Zhang, Xiaodong Cui, Ulrich Finkler, David Kung, and Siyuan Lu. 2020. Map generation from large scale incomplete and inaccurate data labels. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2514–2522.
- [84] Yan Zhang, Yunhuai Liu, Genjian Li, Yi Ding, Ning Chen, Hao Zhang, Tian He, and Desheng Zhang. 2019. Route Prediction for Instant Delivery. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 124 (sep 2019), 25 pages. <https://doi.org/10.1145/3351282>
- [85] Dong Zhao, Zijian Cao, Chen Ju, Desheng Zhang, and Huadong Ma. 2020. D2Park: Diversified Demand-Aware On-Street Parking Guidance. 4, 4, Article 163 (dec 2020), 25 pages. <https://doi.org/10.1145/3432214>
- [86] Lisheng Zhao, Jiali Mao, Min Pu, Guoping Liu, Cheqing Jin, Weining Qian, Aoying Zhou, Xiang Wen, Runbo Hu, and Hua Chai. 2020. Automatic calibration of road intersection topology using trajectories. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*. IEEE, 1633–1644.
- [87] Yan Zhao, Shuo Shang, Yu Wang, Bolong Zheng, Quoc Viet Hung Nguyen, and Kai Zheng. 2018. Rest: A reference-based framework for spatio-temporal trajectory compression. In *KDD'2018*. 2797–2806.
- [88] Pengfei Zhou, Yi Ding, Yang Li, Mo Li, Guobin Shen, and Tian He. 2022. Experience: Adopting Indoor Outdoor Detection in On-demand Food Delivery Business. (2022).
- [89] Pengfei Zhou, Yuanqing Zheng, Zhenjiang Li, Mo Li, and Guobin Shen. 2012. Iodetector: A generic service for indoor outdoor detection. In *Proceedings of the 10th acm conference on embedded network sensor systems*. 113–126.
- [90] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*. 2223–2232.