

Research Article

A Modified Inverse Distance Weighting Method for Interpolation in Open Public Places Based on Wi-Fi Probe Data

Da-wei Wang,¹ Lu-ning Li,¹ Cheng Hu,² Qiang Li ,¹ Xiang Chen ,³ and Pin-wen Huang¹

¹Faculty of Geographical Science, Beijing Normal University, Beijing, China

²Safety and Emergency Management Lab, Beijing Municipal Institute of Labour Protection, Beijing, China

³Department of Geography, University of Connecticut, Storrs, Connecticut, USA

Correspondence should be addressed to Qiang Li; liqiang@bnu.edu.cn

Received 31 January 2019; Revised 15 May 2019; Accepted 12 June 2019; Published 4 July 2019

Guest Editor: Nikolai Bode

Copyright © 2019 Da-wei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Urban open places with a public service function (e.g., urban parks) are likely to be populated in peak hours and during public events. To mitigate the risk of overcrowding and even events of stampedes, it is of considerable significance to realize a real-time full coverage estimate of the population density. The main challenge has been the limited deployment of crowd surveillance detectors in open public spaces, leading to incomplete data coverage and thus impacting the quality and reliability of the density estimation. To remedy this issue, this paper proposes a modified inverse distance weighting (IDW) method, named the inverse distance weighting based on path selection behavior (IDWPSB) method. The proposed IDWPSB method adjusts the distance decay effect according to visitors' path selection behavior, which better characterizes the human dynamics in open spaces. By implementing the model in a real-world road network in the Shichahai scenic area in Beijing, China, the study shows a decrease in the absolute deviation by 17.62% comparing the results between the new method and the traditional IDW method, justifying the effectiveness of the new method for spatial interpolation in open public places. By considering the behavioral factor, the proposed IDWPSB method can provide insights into public safety management with the increasing availability of data derived from location-based services.

1. Introduction

In Chinese cities, open places serving a public service function (e.g., urban parks) are likely to be populated in peak hours and during public events. The inherent difficulties of surveillance and crowd control in these open spaces highlight the numerous evacuation problems that could arise in an emergency, including the occurrence of stampedes. This is precisely what occurred on February 5, 2004, at the Lantern Festival in Beijing's Miyun Park. In this instance, the primary viewing space was trampled and caused 37 fatalities even though the total number of visitors was far below the park's maximum capacity. Hence, controlling the total number of visitors is insufficient for comprehensive safety management in open public places. It is paramount to realize a real-time and full coverage population density estimate to prevent local overcrowding.

To date, many techniques have been applied to crowd surveillance. Video monitoring recognition is one of the

most widely used techniques in public security surveillance. With the development of computer vision, studies on crowd analysis [1, 2] and anomalous detection [3–5] based on video monitoring recognition techniques have made significant strides. Although critical improvements have been made in the video monitoring recognition techniques, it remains infeasible to count the number of people primarily due to the partial occlusions among individuals when the crowd density is extremely high [6]. This difficulty in counting high-density crowds has become a more severe issue for safety management in open public places.

There is an opportunity to overcome this issue by using location-based services (LBSs), such as cellular phone and Wi-Fi probe data. The popularity of smartphones ensures that cell phone data, including cellular signaling data and call detail record data, can fill the gap of individual-based trajectory tracking in urban open spaces [7–10]. Nonetheless, employing cellular phone data has been considered an invasion of people's privacy [11] and, consequently, its widespread

use in exploring individual behavior is restricted. However, the increasing ubiquity of public wireless networks in urban environments creates nascent pathways to understand population dynamics across space and over time [12]. As an emerging location sensor, Wi-Fi probes can conveniently obtain location information from mobile devices [13]. Given the popularity of smartphones today, Wi-Fi probes may prove indispensable in acquiring aggregate movement information in an area of interest [14]. Since Wi-Fi probes merely maintain a unique log of the mobile device, there are little potential privacy infringements [11]. Wi-Fi probe data (called Wi-Fi data hereafter) are consequently becoming an integral component of crowd surveillance [12, 15, 16].

Regardless of how the data collection process ultimately develops, the age-old paradox of expensive detection equipment and limited budgets remains in effect, which leads to the limited deployment of detectors in open public spaces. Moreover, equipment failures and data transmission interruptions diminish the data coverage and severely weaken the quality and reliability of related applications. Thus, the challenge is to realize an acceptable level of the data accuracy within the constraints of a limited budget [17]. The spatial interpolation technique can yield efficient imputations of the missing data while being cost-effective without additional data collection efforts [18]. To this end, various interpolation methods, including the geostatistical and deterministic methods, have been developed in recent decades. The most common geostatistical method applied to traffic interpolation is the Kriging method, which is based on the assumption that the area of interest is an unrestricted region that has spatial autocorrelation with its neighborhoods. Typically, the Kriging method only adopts the Euclidean distance as the distance metric to represent the distance decay effect. Modifications of the Kriging have adopted various distance metrics, such as the road network distance [17, 19] and the great-circle distance [20], which, to a limited degree, can address the limitations associated with the traditional Kriging. The inverse distance weighting (IDW) method is the most widely utilized deterministic method and is commonly applied to big dataset interpolation, including air quality [21] and noise pollution monitoring [22], and has been implemented as a standard spatial interpolation procedure in many geographic information systems (GIS) software packages [23]. The IDW method assumes that each measured point has a local influence on unsampled sites moderated by the distance decay effect. Compared to the Kriging method, the IDW method has a greater ability to adjust the distance metric as it is not restricted by the covariance and variation functions. However, the primary disadvantage of the existing IDW methods is that they only consider objective factors such as the road network structure and spatial distance as decisive factors in the interpolation process while ignoring the influence of individual travel behavior. In an urban open space, an individual's travel often matches their activity demands. The seemingly free-flowing crowd actually has a particular preference for certain paths, referred to as the path selection behavior. The path selection behavior becomes more prominent in emergency situations such as overcrowding [24]. This path selection behavior has

TABLE 1: Description of Wi-Fi data.

Field	Description
Probe ID	Wi-Fi probe label
User MAC	MAC address of the mobile device
Start Timestamp	Start time of receiving a probe request
End Timestamp	End time of receiving a probe request
RSS	Received signal strength of the probe request

not been adequately considered in the existing interpolation methods.

To estimate the spatial distribution of the crowd population in open public places, we have proposed a new interpolation method called the inverse distance weighting method based on path selection behavior (IDWPSB). The path selection behavior is used to adjust the distance decay effect, making the interpolation result more suitable for approximating the actual crowd characteristics. The IDWPSB has been applied to a case study in the Shichahai scenic area in Beijing, China, to verify the computational efficiency and reliability of this new method in estimating the crowd population in road networks.

2. Materials

2.1. Description of Wi-Fi Data. A Wi-Fi-enabled mobile device (e.g., smartphone) can initiate a connection to a Wi-Fi network by continuously broadcasting signals, also known as the probe requests. Each probe request contains a sequence of device information, including the media access control (MAC) address, device type, brand, and manufacturer. For each smart device, the MAC address is unique to the network connection. Since the probe requests are not encrypted, they can be passively captured and decoded with the help of wireless sniffers. In addition, the received signal strength (RSS) of probe requests can also be measured. All these data are uploaded to a server.

The Wi-Fi probe is a type of wireless sniffer. A description of the data collected by the Wi-Fi probe is shown in Table 1. Since there is a clear correspondence between a user and his smart device, the user's MAC address can be regarded as an identifier of a specific individual who is located within the detection range of the Wi-Fi probe. To avoid violations of privacy, the user's MAC address is anonymized by only extracting a fragment from the source record. The timestamp records the temporal information of the Wi-Fi connection. The RSS indicates the connection intensity, which mainly depends on the physical distance between the smart device and the Wi-Fi probe.

2.2. Preprocessing of Wi-Fi Data. Before being analyzed, the raw Wi-Fi data extracted from the server require several steps of preprocessing. First, the RSS needs to be resampled at a regular time step due to the bursty nature of the probe requests, as the probe requests sent by a mobile device are not evenly distributed in time. For example, there can be a sequence of probe requests from one smart device within a few hundred milliseconds, followed by a silent period that

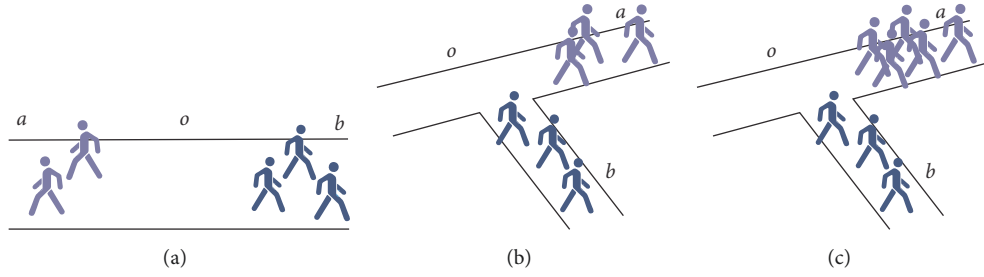


FIGURE 1: Crowd interpolation in different scenarios. Node o is the unsampled site, and nodes a and b are sampling sites.

lasts several seconds. The number of bursts depends on the working condition of the smart device. To mitigate the uncertainty of the RSS captures, we averaged the RSS values received within a one-second interval [25].

The second step is to uniformize the detection radius among the Wi-Fi probes installed in various environments by data screening. The relationship between the RSS and physical distance is calculated as

$$S = 10^{(P-I)/N} \quad (1)$$

where S is the physical distance from the smart device to the Wi-Fi probe, I is the RSS, P is the calibration parameter of the RSS, and N is the calibration parameter of the distance decay effect. For a given type of Wi-Fi probe, P and N are constants. By setting a maximum detection radius for the Wi-Fi probe, the lower limit of the RSS, I_{min} , can be calculated accordingly, by which the request records with RSS values smaller than I_{min} can be excluded. In addition, to remove individuals who are not outdoor, a threshold for the time duration can be set, defined as the end time minus the start time. All the records with a time duration larger than the threshold value (i.e., one hour) are eliminated because individuals in an indoor environment stay connected longer.

Third, considering that some individuals, such as children and senior citizens, do not carry a smart device, we adjusted the detected population through the detectable rate of the Wi-Fi probe. The detectable rate of the Wi-Fi probe is defined as

$$r_d = \frac{M_d}{M_a} \quad (2)$$

where r_d is the detectable rate, M_d is the population detected by the Wi-Fi probe, and M_a is the actual population in the detection radius. The detectable rate of a Wi-Fi probe is related to multiple factors, including the installation environment, working conditions of the smart devices, and the population under coverage. We obtained the parameter in a field survey, in which M_a was acquired by manually counting the number of people within the scope. By surveying the sampling sites, the detectable rate of each Wi-Fi probe is calibrated. Then, the adjusted detected population can be calculated accordingly as

$$M_{ad} = \frac{M_d}{r_d} \quad (3)$$

where M_{ad} is the adjusted population within the detection radius of a Wi-Fi probe.

3. Methodology

3.1. Inverse Distance Weighting (IDW) Method. First, the local crowd density at the sampling site is calculated as

$$\rho_i = \frac{M_i}{S_i} \quad (4)$$

where ρ_i is the local crowd density and M_i is the adjusted detected population at a sampling site i . S_i is the total area of the road segments in the detection radius of Wi-Fi probe i and is calculated as

$$S_i = \sum_{j=1}^n l_j * d_j \quad (5)$$

where n is the number of road segments within the detection radius of Wi-Fi probe i , l_j is the length of road segment j , and d_j is the average width of road segment j .

According to the principle of the IDW, the estimated crowd density at an unsampled site, (node) o , is calculated as

$$\rho_o = \frac{\sum (\rho_i / D_{o,i}^2)}{\sum (1 / D_{o,i}^2)} \quad (6)$$

where ρ_o is the crowd density of the unsampled node, o ; ρ_i is the crowd density of the sampling node, i ; and $D_{o,i}$ is the shortest path distance from node o to node i .

A series of scenarios were set to simulate the interpolation process for open public places and validate the effectiveness of (6). As shown in Figure 1, we assume that node o is the unsampled site and that nodes a and b are sampling sites. We also assume that the path distance from node o to node a and that to node b are equal. Equation (7) formulates the crowd density at node o as ρ_o , which is estimated based on the observed population on nodes a and b .

$$\rho_o = \frac{\rho_a / D_{o,a}^2 + \rho_b / D_{o,b}^2}{1 / D_{o,a}^2 + 1 / D_{o,b}^2} \quad (7)$$

where ρ_a and ρ_b are the observed population on node a and node b , respectively, $D_{o,a}$ and $D_{o,b}$ are the shortest path distances for segment o - a and segment o - b , respectively.

To further explain Equation (7), we have designed three scenarios, as shown in Figure 1. In Figure 1(a), when there are no road branches, $\rho_o = \rho_a + \rho_b$. Now, we add a junction to the

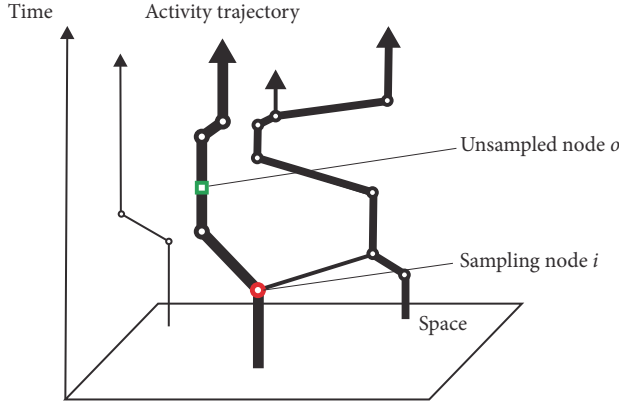


FIGURE 2: Activity trajectory in a three-dimensional space.

road, as shown in Figure 1(b). When individuals select roads randomly (i.e., no path selection behavior), the crowd flow will evenly split between $o-a$ and $o-b$. In this case, node a and node b still contribute equally to the population estimation at node o . In Figure 1(c), when node a becomes more attractive than node b , we adjust the distance weight according to the path selection behavior to reflect the attractiveness of the road segment.

3.2. Inverse Distance Weighting Based on the Path Selection Behavior. Because of the large number of junctions in the road networks of open public places, the distance decay effect with the spatial distance on a straight road is altered by individual preference. Through the observation of individuals' activity trajectories derived from the Wi-Fi data, the path selection behavior pattern, denoted as the transition probability, can be extracted. Adjustment of the distance decay effect according to the transition probability of the path selection behavior can improve the interpolation accuracy.

An activity trajectory Tra_x is defined as a sequence of locations that an individual walks through during a period, denoted as

$$Tra_x = \{(l_j, t_j) \mid j = 1, 2, \dots, n\} \quad (8)$$

where l_j is the j th node on the activity trajectory and t_j is the timestamp of node j , both of which are recorded by a series of Wi-Fi probes located at different sites, and n is the total number of nodes on the activity trajectory. As shown in Figure 2, a cluster of activity trajectories can reveal the diversity of the path selection behavior among the crowd. In a road network, the set of nodes that most activity trajectories contain has a stronger association, regardless of the spatial distance.

The transition probability of the path selection behavior can be calculated from a large number of activity trajectories recorded by Wi-Fi data over a period of time. For a pair of adjacent nodes, $i-o$, the transition probability is defined as the proportion of activity trajectories that contain both nodes i and o to those containing node i , calculated as

$$P_{i,o} = \frac{\sum_1^m Tra_x = \{\dots, (i, t), (o, t+1), \dots\}}{\sum_1^m Tra_x = \{\dots, (i, t'), \dots\}} \quad (9)$$

where m is the total number of individuals recorded by the Wi-Fi probes, and t and t' are arbitrary timestamps on an activity trajectory. For nodes that are not adjacent, the transition probability is the product of the transition probabilities of each pair of adjacent nodes on the shortest path between the unsampled node and the nonadjacent sampling node, calculated as

$$P_{i,o} = \prod_0^k P_{j,j+1} \quad (10)$$

where k is the number of intermediate nodes from node o to node i . It is worth noting that since an activity trajectory has a movement direction, $P_{o,i}$ may not be equal to $P_{i,o}$. Thus, the relationship between node o and node i can be represented as the average of the transition probabilities between nodes o and i , denoted as

$$\bar{P}_{i,o} = \frac{P_{i,o} + P_{o,i}}{2} \quad (11)$$

Then, the transition probability of the path selection behavior is used to adjust the distance decay effect in the traditional IDW metric. The higher the transition probability of the path selection behavior is, the smaller the distance decay is, which is calculated as

$$D_{i,o}^{*2} = \frac{D_{i,o}^2}{\bar{P}_{i,o}} \quad (12)$$

where $D_{i,o}^*$ is the adjusted path distance from node i to node o , $D_{i,o}$ is the shortest path distance, and $\bar{P}_{i,o}$ is the average transition probability between nodes o and i .

3.3. The Validation Index of the Interpolation. Cross-validation is applied to evaluate the accuracy of the interpolation results by comparing the interpolated crowd density of a sampling node with the adjusted crowd density detected by a Wi-Fi probe. As mentioned before, there is a deviation between the population detected by the Wi-Fi probe and the ground truth data because the detectable rate of the Wi-Fi probe is not 100%. For this reason, some field survey experiments were carried out to calibrate the detection rate of the Wi-Fi probe and to adjust the detected population according to Equation (3), making the adjusted population as close to the ground truth value as possible. Although there is still a deviation between the adjusted population and the ground truth data, assuming that this deviation is a systematic error for each Wi-Fi probe, it is acceptable to use the adjusted population as the ground truth because it does not affect the validation result.

The widely used accuracy indices, the mean absolute error (MAE), and mean bias error (MBE) are used to indicate the absolute deviation and bias of the interpolation results, calculated as

$$MAE = \frac{\sum_{i=1}^n |\rho_{adj,i} - \rho_{est,i}|}{n} \quad (13)$$

$$MBE = \frac{\sum_{i=1}^n (\rho_{adj,i} - \rho_{est,i})}{n} \quad (14)$$



FIGURE 3: The locations of the installed Wi-Fi probes with the crowd density on a (a) weekday and (b) a holiday.

where $\rho_{adj,i}$ is the adjusted crowd density detected by a Wi-Fi probe on node i , which is regarded as the true value; $\rho_{est,i}$ is the estimated value obtained by IDWPSB method on node i .

The MAE represents the mean value of the absolute deviations between the estimated values and the true values. The smaller the MAE is, the closer the estimated values are to the true values. The MBE denotes the mean bias between the estimated values and the true values; the sign of the MBE suggests underestimation (positive value) or overestimation (negative) of the interpolation results.

4. Study Area and Data Collection

The Shichahai scenic area is a tourist neighborhood in Northwest Beijing, covering an area of 2.31 million square meters. As an open public place in a metropolitan center, the Shichahai scenic area fulfills multiple public service functions, combining a tourist attraction with business and residential sectors. Because of the multifunction role of the area, it is crowded throughout the year. On a typical holiday, the crowd flow can easily reach 20 thousand per hour. Thus, there has been a considerable safety concern and needs of crowd control in this area.

As shown in Figure 3, the road network in this area is rather complex and contains many junctions. To select

the optimal locations to install the Wi-Fi probes, first, the road segments considered to be crowded were designated by experts, as illustrated by the purple lines in Figure 3. Then, the three tourist landmarks, including the Yinding Bridge, the former residence of Soong Ching Ling, and the Prince Gong's Mansion, illustrated by pink hollow stars in Figure 3, were also selected as candidates for installation. Taking these factors into account, 23 Wi-Fi probes, denoted as $p_1, p_2 \dots p_{23}$, were installed along the main roads to collect the crowd data. By preprocessing the Wi-Fi data, the detection radius of the Wi-Fi probe was uniformized to 50 m.

During the study period from December 23, 2016, to March 16, 2017, the crowd data of more than 6,000,000 individuals were collected by the installed Wi-Fi probes. The mean value of the all-day crowd density during the study period was calculated according to Equation (3) for each sampling site and is shown in Figures 3(a) and 3(b) for a weekday and a holiday. Table 2 shows the exact values of the crowd density at each sampling site and ranks them by sorting the weekday and holiday values in descending order. It can be seen from the table that although the absolute crowd density value differs by the two days, there is no significant difference in the ranks of the crowd density. That is, whether on the weekday or the holiday, the crowding patterns are similar in

TABLE 2: The summary of the crowd density on a holiday and a weekday.

Probe ID	Crowd density on a weekday (people per square meter)	Rank of crowd density on a weekday	Crowd density on a holiday (people per square meter)	Rank of crowd density on a holiday
1	0.124	16	0.290	16
2	0.303	8	0.629	8
3	0.016	22	0.137	19
4	0.218	10	0.473	10
5	0.199	11	0.418	11
6	0.151	15	0.331	15
7	0.325	7	0.702	7
8	0.901	1	1.895	1
9	0.451	5	0.949	5
10	0.033	21	0.063	22
11	0.804	2	1.636	2
12	0.619	4	1.341	4
13	0.156	13	0.342	14
14	0.662	3	1.357	3
15	0.106	17	0.233	17
16	0.008	23	0.008	23
17	0.079	18	0.172	18
18	0.046	20	0.107	21
19	0.055	19	0.109	20
20	0.181	12	0.387	12
21	0.429	6	0.912	6
22	0.250	9	0.545	9
23	0.153	14	0.343	13

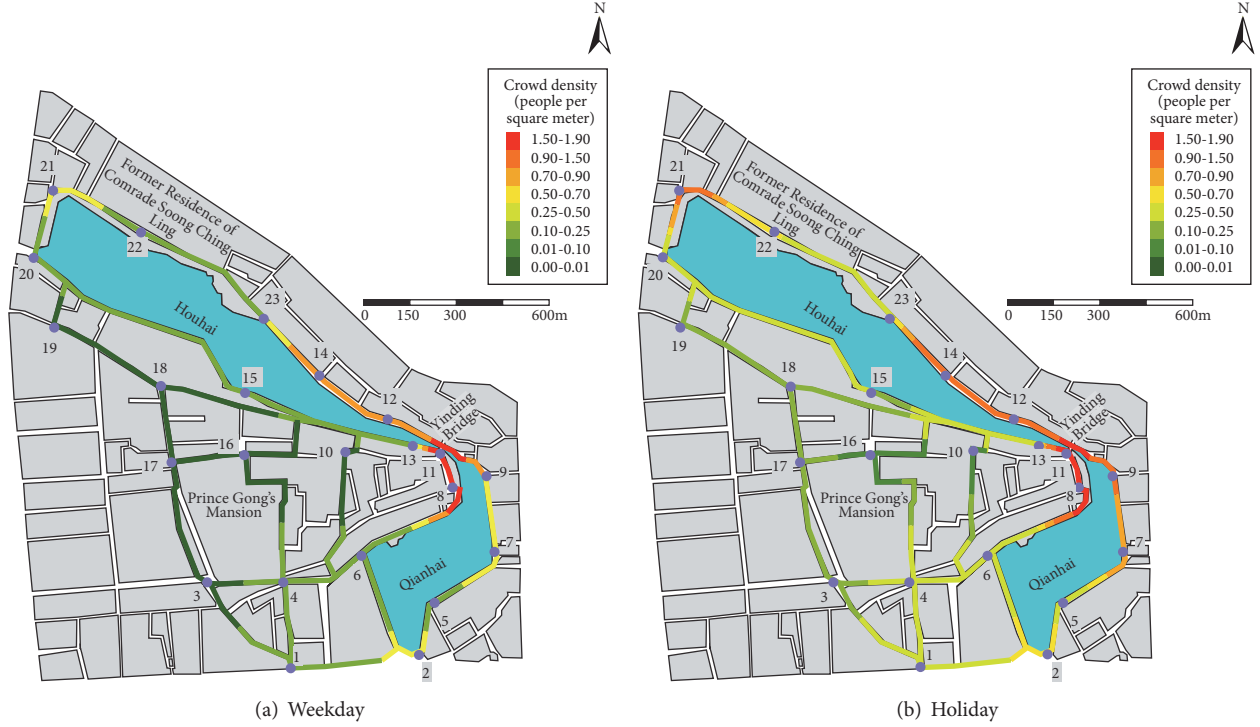


FIGURE 4: The estimate of the global crowd density by IDWPSB method in the study area on (a) weekday and (b) a holiday.

the study area. The most crowded sites are near the Yinding Bridge, including p8, p11, p14, and p12, which are the core scenic spots.

5. Results and Discussion

5.1. Comparing the Global Crowd Density Estimates by the IDWPSB and IDW Methods. As shown in Figure 4, we selected a typical peak time (3 p.m.) on a weekday (Feb 8, 2017) and a holiday (Dec 25, 2016) as an example to show the interpolation results of the IDWPSB method. The spatial resolution of the interpolation was 5 meter, generating the estimation of 628 unsampled sites in total. Overall, the global crowd density on the holiday was significantly higher than that on the weekday, with more road segments highly populated. The crowd tended to aggregate towards the northeast corner, the Yinding Bridge, on both days; meanwhile, the population was relatively sparse in the southwest quadrant, which is the residential sector.

Additionally, the traditional IDW method was employed to interpolate the population density and was compared with the new IDWPSB method. As shown in Figure 5, the majority of the MAE values at the sampling sites obtained by the IDWPSB method are smaller than those obtained by the IDW method. As shown by the MBE values, the IDWPSB method is less likely to overestimate or underestimate the crowd density in the cross-validation, suggesting a more robust performance than the IDW method. Furthermore, the accuracy improvement ratio, P_{imp} , was calculated as the

difference between the IDWPSB MAE and the IDW MAE, as shown in

$$P_{imp} = \frac{\sum MAE_{IDWPSB} - \sum MAE_{IDW}}{\sum MAE_{IDW}} \quad (15)$$

For the weekday result, the accuracy improvement ratio is 20.49% and for the holiday, the ratio is 15.69%. In general, the IDWPSB method improves the interpolation accuracy by 17.62%, indicating a better performance than the IDW. In terms of computational efficiency, the interpolation process took 52.4 seconds using the IDW method and 74.3 seconds using the IDWPSB method under the same computing environment. These various aspects of comparison suggest that including the path selection behavior in the IDWPSB method can improve the interpolation accuracy without significantly increasing the computational demand.

5.2. Factors Influencing the Interpolation Accuracy. Several factors may affect the interpolation accuracy and should be taken into consideration. The first factor is the number of sampling sites. We implemented the IDWPSB interpolation by using different numbers of sampling sites ranging from 10 to 22 and then randomly selecting sampling sites from the pool. Figure 6 shows the variation in the MAE value for each sampling site when the site is estimated using cross-validation. It is evident that there is a decreasing trend in the MAE when the number of sampling sites increases in all instances. In reality, the optimal number of sampling sites should be determined by both budget consideration and requirements for accuracy.

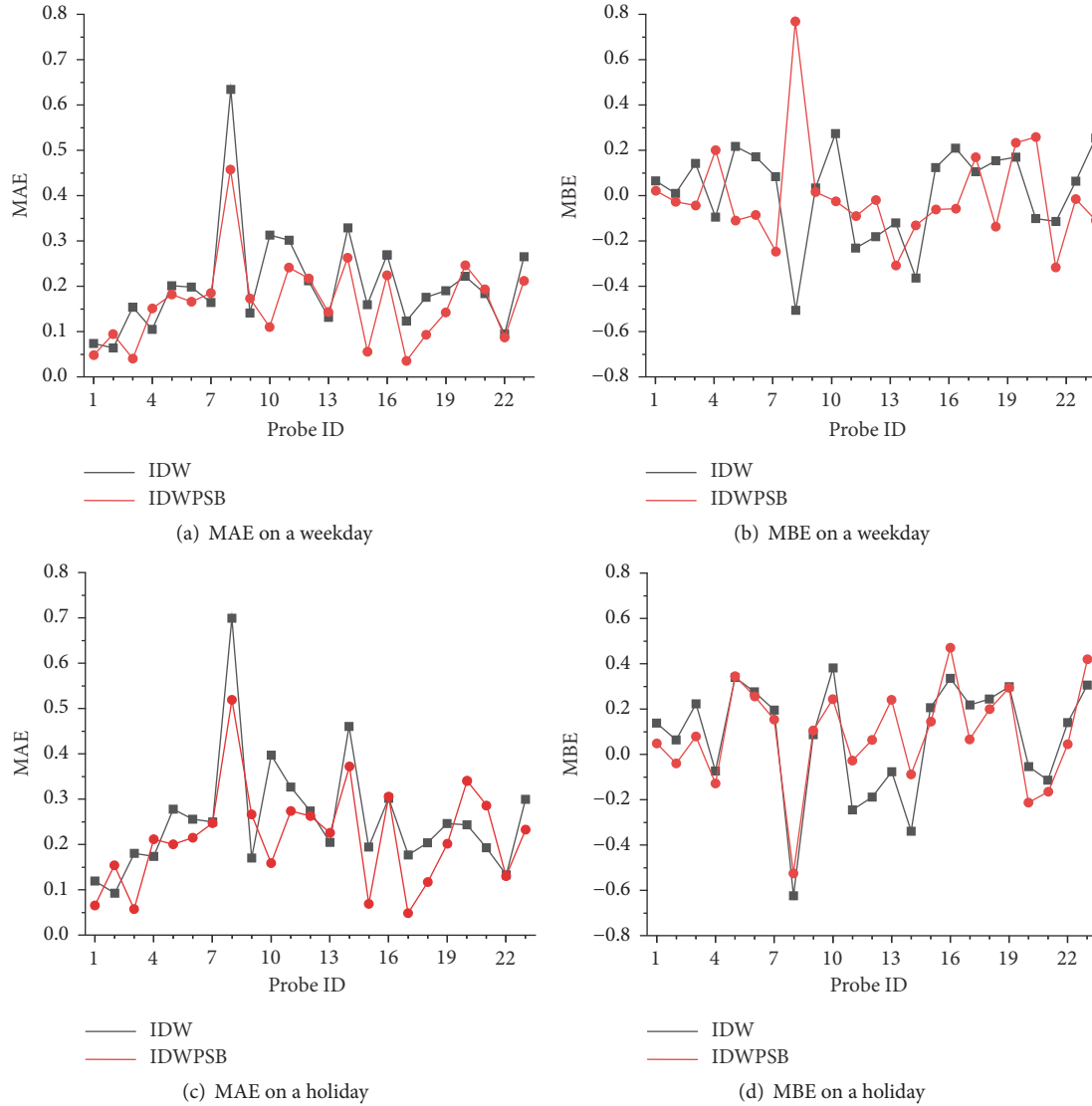


FIGURE 5: Comparison of the interpolation results in MAE and MBE between the IDW and IDWPSB methods.

The second factor is the location of the sampling sites. To evaluate the relationship of the interpolation error with respect to the location of the sampling sites, the all-day crowd density for each sampling site on the weekday and the holiday was estimated for all the other sampling sites using the IDWPSB method. The spatial distributions of the MAE and MBE are shown in Figure 7. It is worth noting that, on either day, the site with the maximum interpolation error is p8, which is the sampling site with the highest crowd density. This result is likely due to the inherent bias of the IDW method in that it fails to extrapolate the maximum value. Thus, we suggest that the deployment of the Wi-Fi probes should cover the sites with the local maximum values.

The third factor is the role of an unsampled node in the road network. We use the node centrality to indicate the importance of an unsampled node in the road network. Here, the centrality of node i is defined as the number of times that node i is on the shortest path between any other two

nodes being the origin and the destination. As shown in Figure 8, we investigated the relationship between the node centrality of an unsampled node and its interpolation MAE in the cross-validation. It is evident that there is no relationship between the node centrality and the MAE, suggesting that the accuracy of the interpolation is hardly affected by the role of an unsampled node in the road network. An advantage of the new method is that it requires less consideration of the road network structure.

6. Conclusion

A real-time full coverage estimate of crowd density is essential for safety management of urban spaces. In practice, the trade-off between full coverage monitoring and limited budgets always exists, leading to the partial deployment of crowd surveillance detectors in open public spaces. The incomplete data coverage seriously influences the quality and reliability of

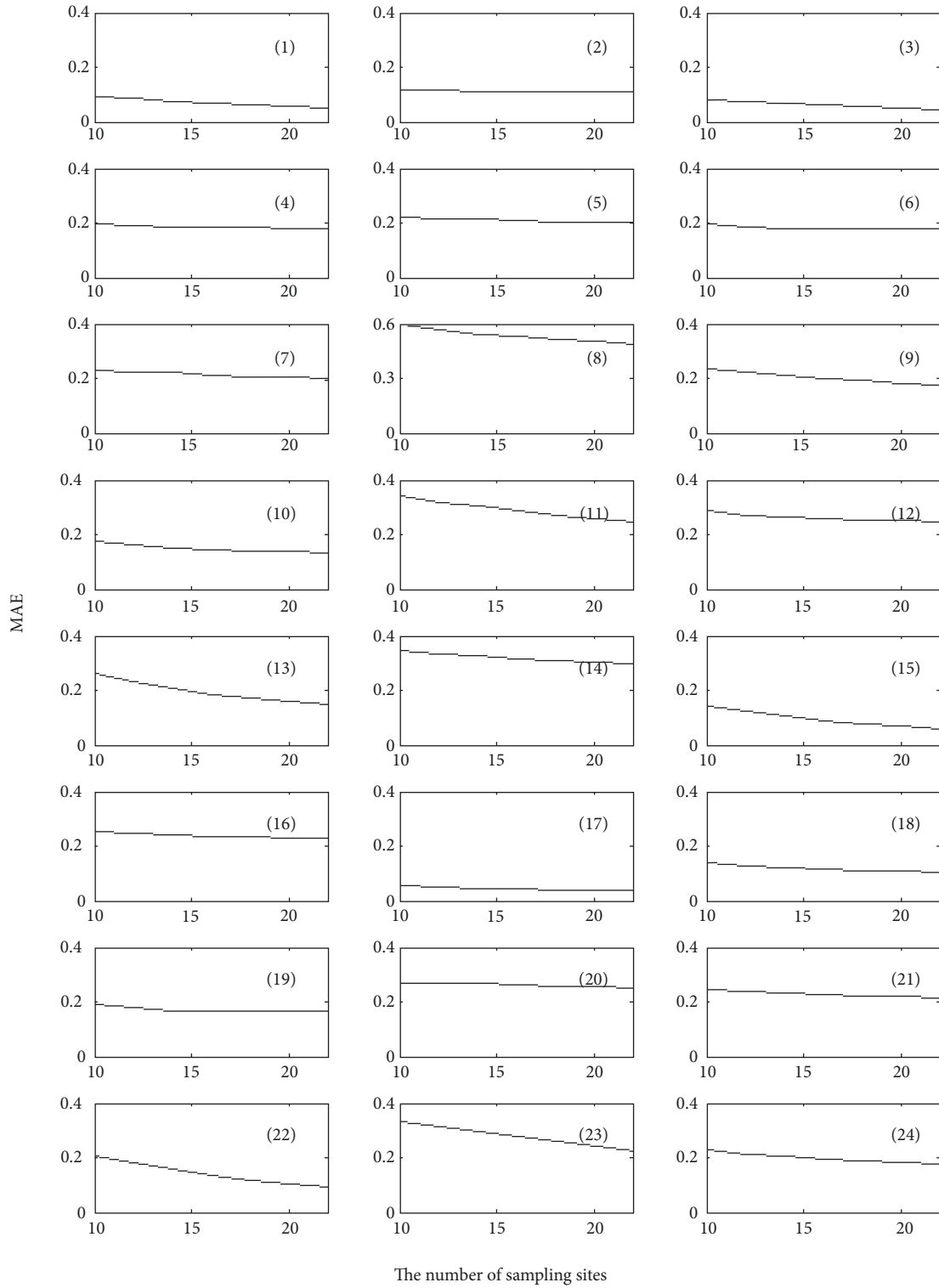


FIGURE 6: Change of MAE by the number of sampling sites: (1)–(23) corresponding to sampling site p1–p23 when the site is selected in cross-validation and (24) is the overall MAE.

the crowd density estimate. By extracting individual activity trajectories from the Wi-Fi data, this paper proposes a modified IDW method based on the path selection behaviors, named the IDWPSB method. The main improvement of the

new method is to adjust the distance decay effect in the IDW method according to the transition probabilities of the path selection behavior. The case study in the Shichahai scenic area of Beijing city demonstrates better performance

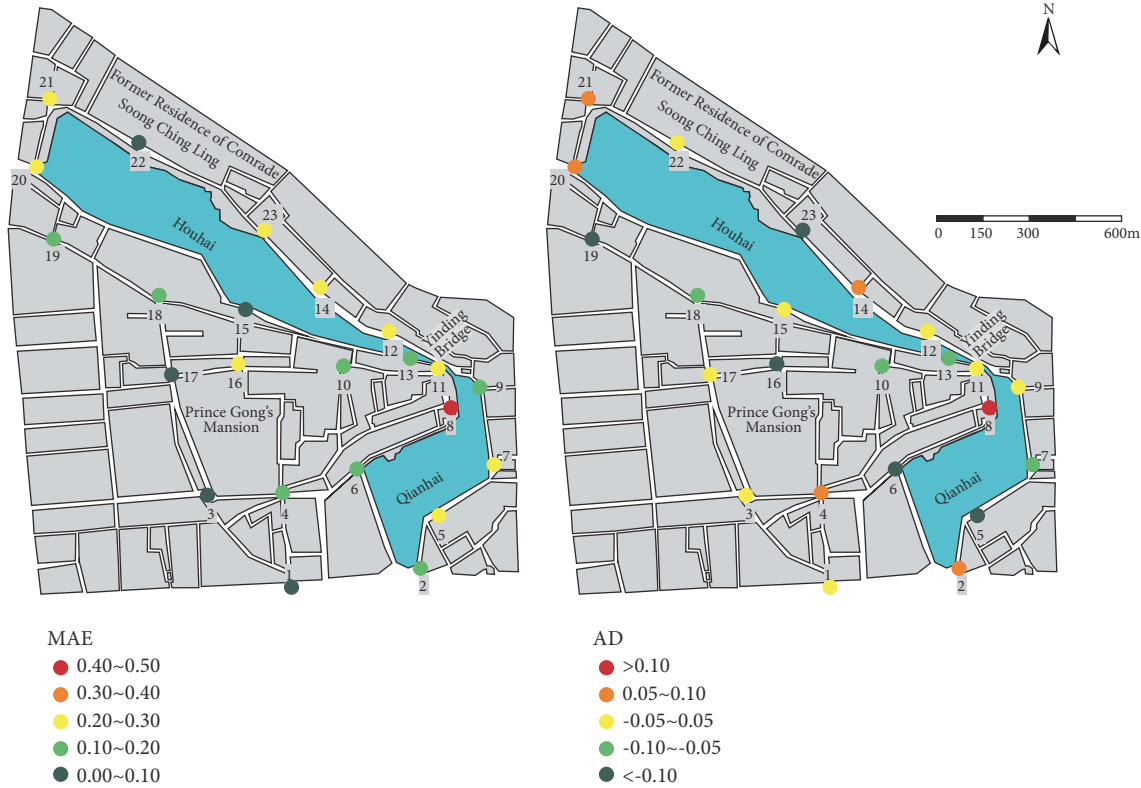


FIGURE 7: The spatial distribution of the interpolation errors (MAE and MBE) in the cross-validation.

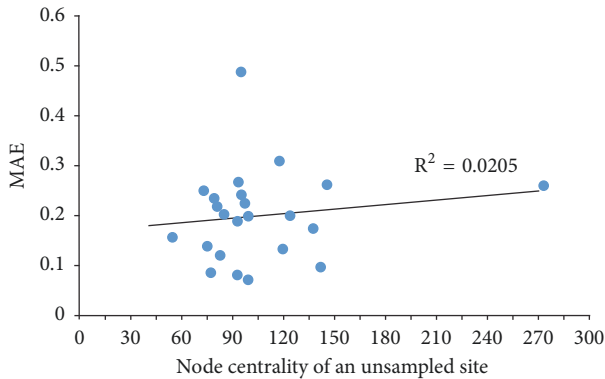


FIGURE 8: The relationship of the MAE for the node centrality of an unsampled site.

of the proposed method than the traditional IDW method. The proposed method can also be applied to urban spaces structured by networks in both indoor and outdoor environments with the availability of LBS data. However, it should be noted that the path preference is not static in time. Thus, the transition probabilities of the path selection behavior should be a function of time, which may vary over time of the day or may differ between weekdays and weekends. The time-dependent calculation will provide a valid context for improving the interpolation method and accommodating dynamic travel environments.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest in the research.

Acknowledgments

This work was supported by the National Key Research and Development Program of China 2017YFC1503004.

References

- [1] C. Lijun and H. Kaiqi, "Video-based crowd density estimation and prediction system for wide-area surveillance," *China Communications*, vol. 10, no. 5, pp. 79–88, 2013.
- [2] H. Fradi and J. Dugelay, "Towards crowd density-aware video surveillance applications," *Information Fusion*, vol. 24, pp. 3–15, 2015.
- [3] D.-Y. Chen and P.-C. Huang, "Dynamic human crowd modeling and its application to anomalous events detection," in *Proceedings of the IEEE International Conference on Multimedia and Expo, ICME '10*, pp. 1582–1587, 2010.
- [4] F. Jiang, Y. Wu, and A. K. Katsaggelos, "Detecting contextual anomalies of crowd motion in surveillance video," in *Proceedings of the 16th IEEE International Conference on Image Processing ICIP '09*, pp. 1113–1116, 2009.

- [5] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18–32, 2014.
- [6] V. Eiselein, H. Fradi, I. Keller, T. Sikora, and J.-L. Dugelay, "Enhancing human detection using crowd density measures and an adaptive correction filter," in *Proceedings of the 10th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS '13*, pp. 19–24, 2013.
- [7] Y. Asakura and E. Hato, "Tracking survey for individual travel behaviour using mobile communication instruments," *Transportation Research Part C: Emerging Technologies*, vol. 12, no. 3–4, pp. 273–291, 2004.
- [8] F. Calabrese, M. Colonna, P. Lovisolo, D. Parata, and C. Ratti, "Real-time urban monitoring using cell phones: a case study in Rome," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 141–151, 2011.
- [9] F. Calabrese, G. Di Lorenzo, L. Liu, and C. Ratti, "Estimating origin-destination flows using mobile phone location data," *IEEE Pervasive Computing*, vol. 10, no. 4, pp. 36–44, 2011.
- [10] S. Jiang, G. A. Fiore, Y. Yang, J. Ferreira Jr., E. Frazzoli, and M. C. González, "A review of urban computing for mobile phone traces: Current methods, challenges and opportunities," in *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, UrbComp 2013 - Held in Conjunction with the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '13*, vol. 1, ACM, New York, NY, USA, 2013.
- [11] N. Abedi, A. Bhaskar, and E. Chung, "Tracking spatio-temporal movement of human in terms of space utilization using media-access-control address data," *Applied Geography*, vol. 51, pp. 72–81, 2014.
- [12] C. E. Kontokosta and N. Johnson, "Urban phenology: toward a real-time census of the city using Wi-Fi data," *Computers, Environment and Urban Systems*, vol. 64, pp. 144–153, 2017.
- [13] M. Cunche, "I know your MAC address: targeted tracking of individual using Wi-Fi," *Journal of Computer Virology and Hacking Techniques*, vol. 10, no. 4, pp. 219–227, 2014.
- [14] A. B. M. Musa and J. Eriksson, "Tracking unmodified smart-phones using wi-fi monitors," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems (SenSys '12)*, pp. 281–294, ACM, Ontario, Canada, 2012.
- [15] F. Wang, X. Zhu, and J. Miao, "Semantic trajectories-based social relationships discovery using WiFi monitors," *Personal and Ubiquitous Computing*, vol. 21, no. 1, pp. 85–96, 2017.
- [16] Z. Xu, K. Sandrasegaran, X. Kong et al., "Pedestrian monitoring system using Wi-Fi technology and RSSI based localization," *International Journal of Wireless & Mobile Networks*, vol. 5, pp. 17–34, 2013.
- [17] H. Zou, Y. Yue, Q. Li, and A. G. Yeh, "An improved distance metric for the interpolation of link-based traffic data using kriging: a case study of a large-scale urban road network," *International Journal of Geographical Information Science*, vol. 26, no. 4, pp. 667–689, 2012.
- [18] T. J. Klatko, T. U. Saeed, M. Volovski, S. Labi, J. D. Fricker, and K. C. Sinha, "Addressing the local road VMT estimation problem using spatial interpolation techniques," *Journal of Transportation Engineering Part A: Systems*, vol. 143, 2017.
- [19] J. Gutiérrez and J. C. García-Palomares, "Distance-measure impacts on the calculation of transport service areas using GIS," *Environment and Planning B: Planning and Design*, vol. 35, no. 3, pp. 480–503, 2008.
- [20] H. Miura, "A study of travel time prediction using universal kriging," *TOP*, vol. 18, no. 1, pp. 257–270, 2010.
- [21] L. Li, T. Lossner, C. Yorke, and R. Piltner, "Fast inverse distance weighting-based spatiotemporal interpolation: a web-based application of interpolating daily fine particulate matter PM_{2.5} in the contiguous U.S. using parallel programming and k-d tree," *International Journal of Environmental Research and Public Health*, vol. 11, no. 9, pp. 9101–9141, 2014.
- [22] B. I. Harman, H. Koseoglu, and C. O. Yigit, "Performance evaluation of IDW, Kriging and multiquadric interpolation methods in producing noise mapping: a case study at the city of Isparta, Turkey," *Applied Acoustics*, vol. 112, pp. 147–157, 2016.
- [23] W. Luo, M. C. Taylor, and S. R. Parker, "A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales," *International Journal of Climatology*, vol. 28, no. 7, pp. 947–959, 2008.
- [24] X. Chen, M.-P. Kwan, Q. Li, and J. Chen, "A model for evacuation risk assessment with consideration of pre- and post-disaster factors," *Computers, Environment and Urban Systems*, vol. 36, no. 3, pp. 207–217, 2012.
- [25] E. Vattapparamban, B. S. Çiftler, I. Güvenç, K. Akkaya, and A. Kadri, "Indoor occupancy tracking in smart buildings using passive sniffing of probe requests," in *Proceedings of the IEEE International Conference on Communications Workshops, ICC '16*, pp. 38–44, 2016.

