

Received January 19, 2020, accepted February 6, 2020, date of publication February 11, 2020, date of current version March 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2973177

# Spatiotemporal Patterns of Visitors in Urban Green Parks by Mining Social Media Big Data Based Upon WHO Reports

HIDAYAT ULLAH<sup>ID</sup>, WANGGEN WAN<sup>ID</sup>, (Senior Member, IEEE), SAQIB ALI HAIDERY<sup>ID</sup>,  
NAIMAT ULLAH KHAN<sup>ID</sup>, ZEINAB EBRAHIMPOUR<sup>ID</sup>, AND A. A. M. MUZAHID<sup>ID</sup>

School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China  
Institute of Smart City, Shanghai University, Shanghai 200444, China

Corresponding author: Hidayat Ullah (hidayat@shu.edu.cn)

This work was supported by the Anhui Natural Science Foundation (No: 1908085MF178) Anhui Excellent Young Talents Support Program Project (No: gxyqZD2019069) and a project of Shanghai Science and Technology Commission (No: 18510760300).

**ABSTRACT** Green parks in urban areas are believed to enhance the well-being of residents. The importance of green spaces to support health and fitness in urban areas has recently regained interest. Reports released in 2010–2016 by the World Health Organization (WHO) on urban planning, environment, and health stated that green spaces can have a positive impact on physical activity, social and mental well-being, enhance air quality and decrease noise exposure. We analyzed the number of check-ins in various parks of Shanghai by utilizing geotagged social media network check-in data. This article presents a descriptive study using social media data by obtaining the three-year comparison of spatial and temporal patterns of park visits to raise public awareness that green parks provide a healthy environment that can be beneficial for the well-being of urban citizens. We investigated the visitor spatiotemporal behavior in more than 115 green parks in 10 districts of Shanghai with approximately 250,000 check-ins. We examined 3 years of geotagged data and our main findings are: (i) the spatial and temporal variations of users in urban green parks (ii) the gender differences in space and time with relation to urban green parks. The main objective of this article is to present evident data for policymakers on the advantages of providing green spaces access to urban citizens and to facilitate cities with systematic approaches to provide green space access to improve the health of urban citizens.

**INDEX TERMS** Urban green parks, big data, social networks, spatiotemporal, KDE, data mining.

## I. INTRODUCTION

Urban green spaces, especially in parks, provide a wide range of advantages, such as mental and physical health and social benefits [1]. Park visits provide opportunities to explore the advantages of a natural environment especially for citizens living with limited natural interaction [2], [3]. Quantifying visits to urban green parks by residents and to understand the factors influencing their visits are crucial for urban park planning and management. Traditional techniques for analyzing visitor numbers include visitor surveys, direct observation, interviews and on-site counters [4], [5]. Such techniques of systematic observation usually choose or select a representative sample of urban green parks and gather data on park utilization, such as the frequency of visits, activities, and behavior. However, such techniques are typically

site-specific as well as inefficient and therefore have limited spatial coverage [6], [7].

Urban areas are usually characterized by degraded environments, growing pollution, increased temperatures, and declining urban green spaces. Urban green spaces are essential for enhancing the quality of life in urban areas, balancing the heat budget and offering thermal comfort [8] and assisting people to recover from the mental and physical stress of their day-to-day lives [9]. Urban green spaces are essential to improve the living circumstances in urban areas by enhancing air quality and aesthetics, which ultimately results in increasing real estate values and lowering energy usage for cooling. Urban green spaces can also be used as an extensive tool for viable urban areas development [10]. Urban green spaces also provide much-needed space for kids to play, which is essential for their social, physical and cognitive development [11]. Open public spaces were established in the UK (United Kingdom) and the USA (United States of

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang<sup>ID</sup>.

America) in the 19th century to improve the health and quality of life of the working classes who were living in miserable conditions [12]. This study showed that the distance from home to the parks also has an effect on the frequency of the green parks usage when it is first designed; people have confidence in the possible health impacts that may result from open green spaces. They expressed the hope that the green parks would decrease diseases, social conflict and crime, as well as provide “green lungs” for metropolitan places [12]. Easy access to urban green parks is of growing importance in the design of sustainable, safe and viable cities [13], [14]. The researchers neglect some spatial aspects (mostly when performing sentiment analysis using Twitter data) [13], [15], or, on the contrary, studies that consider only the spatial aspects of green park visits [15], [16] and the improvement of health [17]. Even just a limited number of studies have concentrated on the problem of the direct influence of park visits on physical activities or health [18], [19].

Although studies to date have enhanced our knowledge of factors related to psychological disorders, a major restriction of previous research is that it depends strongly on small, often homogeneous samples of people who might not be normally representative of the larger population. Furthermore, these investigations are typically focused on surveys, dependent on retrospective self-reports about mood and health observations: a technique that restricts temporal granularity. In other words, such assessments are intended to compile high-level summaries of experience over long periods.

Many studies have shown an awareness of spatial physical factors affecting park access and use [20], [21]. The most significant factors influencing park use were found to be the natural environment, the sense of space, the level of quietness, and the accessible park services, while the authors stated that distance to the park was the main factor that determines park accessibility [22]. Several other studies highlighted that socio-demographic factors could be offset to the significance of such spatial-physical factors [23], [24]. For example, Byrne et al., 2009 [25] observed that cultural reasons cannot fascinate people to visit neighboring parks. To fill this knowledge gap, both physical and non-physical dimensions have been examined in recent studies [26], [27]. For example, Wang et al. [28] mentioned that both social and physical factors, such as green park closeness and a pleasant walking environment, were statistically important to the perceived accessibility to the parks in Brisbane, Australia. The use of the park facilities is also linked to the accessibility of means of transport and their actual use among residents. However, these factors have also been studied individually or thoroughly in previous research, and the relationship between these factors and park facilities might still be varied based upon the location of green parks and city context.

Kernel density estimation (KDE) is a method to consider disclosure at any point within a spatial frame, irrespective of administrative boundaries. A detailed comparison between two similarly populated areas was carried out using KDE to study the disparities in health or any other issues related to

public health [29]. Earlier, KDE was used as a Geographic Information System (GIS) method to measure the effects of neighborhood features on chronic disease rates, including coronary heart disease and other stress-mediated, dietary and weak-developing contributing agents [29], [30]. KDE could be used to fulfill the frequent need to analyze health data from fields of different sizes and populations [29], [31]. Authors in [32] used social media data and KDE to examine the spatial and temporal patterns of use of urban green parks and the external factors that contribute to their use. The density of food shops was measured differently in order to assess the availability or accessibility of food stores or fast food outlets [33], [34]. Frequently used measures include, for example, basic density techniques, i.e., the number per area [35], [36] or capita [37], or kernel density techniques that give more weight to restaurants and stores, depending upon the distance to the point of observation [38].

Reports from the WHO have previously provided evidence and recommendations on access to green spaces in regard to public health advantages. The reports of WHO based upon urban planning, the environment, and health released in 2010 [39] reveals that urban green spaces can have a positive impact on physical activity, psychological and social well-being, enhance air quality and reduce noise exposure. Another WHO report examined the impact of green spaces on physical activity and their ability to reduce inequalities in public health. This report states that “access to public open space and green areas with appropriate recreation facilities for all age groups is needed to support active recreation”; however, it identifies that intersectoral and multidisciplinary approaches might be required to support disadvantaged groups with the lowest levels of physical activity [40].

Different models have been implemented to define the connection between green spaces and health, Hartig, T. et al., 2014 [41] recommended four major principles and interactive structures from which nature or green spaces may contribute to health, enhanced physical activity, better air quality, reduced stress and increased social cohesion. Numerous studies have revealed that green spaces, for example, public green parks, can decrease noise and air pollution [42], [43], increase mood and related mental outcomes [44], positively affect self-reported health [45], [46], lower cumulative danger of cardiometabolic diseases [47] and reduced metabolic syndrome scores [48]. There is also some evidence that walking in a natural atmosphere compared to urban areas has benefits in terms of physical and psychological rehabilitation in young people [49] and also in elderly patients with high blood pressure [50].

In this study, our objective is to examine the behavior of the public toward the green parks visits using Weibo as a source of a location-based social network (LBSN). The contributions of the paper are twofold, and can be summarized as follows: (i) The difference in user’s behavior towards green parks over 3 years, either by the impact of public awareness reports released by WHO or more interest in using social networks

while performing activities in green parks and, (ii) Gender preference to utilize the green parks facilities for the purpose of health restoration or interest in using location-based social networks.

## II. MATERIALS AND METHODS

In this article, we examined the spatiotemporal park visiting patterns of 115 parks along with 250,632 check-ins. The data is collected from the famous Chinese microblog sina-weibo, also known as weibo, and our main findings are as follows: Comparison of the three-year check-in data collected from weibo for the spatial analysis yearly, weekdays and week-ends. Also, we observed the monthly check-in behavior of the people towards green spaces that represent real-life trends. We also investigated gender differences through spatial and temporal patterns. This strategy of capturing visitor behavior in urban green parks is an attempt to address the shortcomings of earlier studies. This approach paves the way for the use of crowded and social network data in the socio-ecological field research instead of depending on the results of observational data and subjective reporting.

### A. STUDY AREA AND DATASET

The metropolitan city of Shanghai along with the Yangtze River Delta is a part of the alluvial plain, with an average altitude of about 4 m. From east to west, there are slight topographical slopes. The topography is a flat plain except for a few foothills in the southwest. Shanghai belongs to the northern subtropical humid monsoon climatic area with four seasons, full sunshine and plenty of rainfall. Therefore, hot extremes and urbanization in such a metropolitan area have a significant impact on both the public's health and the state's economy. Shanghai is the tenth-largest and most important agglomeration area in the globe and also claims a process of urbanization, which is both relatively early and extremely rapid. When the UN (United Nations) revealed the future forecast of prospects for urbanization, it reported that the urban population of Shanghai was ranked second in the world and first in China [51]. Shanghai City is one of the fastest-growing major cities in the globe with 22,125,000 people living in an area of 4,015 [52]. In the Shanghai municipal boundaries, the total number of green spaces is 366 [53]. The study area includes the 10 districts with green parks, which makes it easier for the residents of the city to engage in healthy activities.

The city of Shanghai was divided into 16 county-level divisions in 2016: 15 districts (Baoshan, Changning, Fengxian, Hongkou, Huangpu, Jiading, Jingan, Jinshan, Minhang, Pudong New Area, Putuo, Qingpu, Songjiang, Xuhui, and Yangpu) and 1 county (Chongming) [54]. Seven districts (Changning, Hongkou, Huangpu, Jingan, Putuo, Xuhui and Yangpu) are situated in Puxi (literally Huangpu West). These seven districts are called the city center or downtown of Shanghai [55]. The study area considered for this study consists of 10 districts, including Baoshan, Changning,



FIGURE 1. Study area.

Hongkou, Huangpu, Jingan, Minhang, Pudong New Area, Putuo, Xuhui, and Yangpu, as shown in Figure 1.

The dataset used to access the number of urban green park visits is gathered from a popular Chinese micro-blog “Weibo”. Weibo is indeed the Chinese Twitter equivalent as well as the largest social media blog in China. The number of Weibo users is very large and it represents the largest available geo-tagged datasets. Based on the most recent Weibo's yearly report, Weibo declared that in 2018 they had more than 500 million registered users actively using the platform, and in December 2018 they had monthly active users hitting 462 million [56]. The demographic characteristics of the users of Weibo are indeed not consistent with the overall population.

Public interfaces for Weibo's location-based services (LBS) were launched on 28 May 2012. Since then, Weibo users have been able to share their real-time locations on the Internet. As one type of fine-scaled accessible crowdsourcing data, the Weibo check-in data was the most suitable dataset we could acquire as a proxy for estimating real park visits. However, earlier research concentrating on 87 urban parks in Shanghai, China, has shown an important connection between Jiebang social media platform check-in data and official visitation statistics [57]. Furthermore, previous research has shown that data from Weibo check-ins can well represent the preferences and activities of people in urban areas.

Utilizing Weibo check-in data [58] a novel model has been suggested that integrates various dimensions of the urban

environment function network and thus enriches the urban network system description. Moreover, while using Weibo check-ins as a visitation proxy is still rare, previous research using data from comparable social media platforms, such as Twitter, Instagram, and Flickr, have observed important beneficial associations between official visitor stats and the number of visitors estimated on these platforms [7], [59]. Weibo application program interfaces (APIs) facilitates the provision for the collection of data.

The dataset was retrieved from Weibo and covered a period of 3 years, from July 2014 to June 2017. After accessing this information about the location, we came to find that some of the locations we got were not green parks (for example, pedestrian walkways, the former residences of celebrities and sculptures). We examined the locations one by one and deleted those that were not within the park category but we included the parking areas that were just connected or a part of a park. Some parks that are bigger in the area and have more than two location IDs, such as garden areas, children playing areas and BBQ areas have been merged into a single location ID. After pre-processing and cleaning, a total of 250632 geo-tagged visits to parks in 10 districts of Shanghai were considered for the study. Data collection was done using the Python Programming Language (version 2.7.12). The data was filtered for anomalies, fake users and invalid records, which includes;

- The geographical location of data resides only in Shanghai;
- The minimum number of check-ins per park must be 100 within the duration of the study;
- Each record must have a user ID, geo-location (latitude and longitude), time, day, month, year and gender; and
- Parks that are divided into multiple geo-locations within the green parks were merged into a single geo-location.

## B. METHODS

In the current study, we analyzed the Weibo based geo-location dataset (July 2014 to June 2017) in 10 districts of Shanghai, China. Figure 2 presents the check-in behavior analytics framework, where the LBSN data analysis methodology involves; data pre-processing and cleaning, spatial and temporal analysis of LBSN data and statistical evaluation to provide the significance of LBSN data.

### 1) DATA PREPARATION

The data collection includes all the check-ins made within the Shanghai boundaries from July 2014 to June 2017. The downloaded data was available in multiple JSON files. Figure 3 reflects the process flow of data preparation;

The Weibo dataset that is used for the current study contains information such as the unique user ID and time and date of check-in. Additionally, the information about geo-location (longitude and latitude) and gender were collected via the Weibo API. Therefore, it is assumed that the LBSN dataset archives day-to-day patterns of activity, behaviors of use

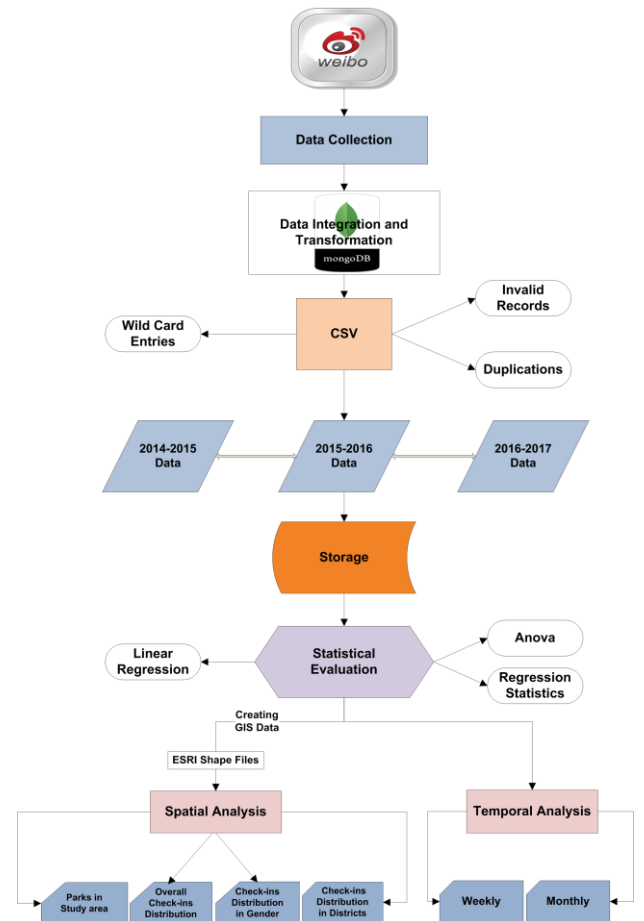


FIGURE 2. Methodology.



FIGURE 3. Process flow of data preparation.

towards social media, and spatial-temporal evidence, which are related to the daily routines of users [60]. JSON is a Java programming platform format that is the most widely used data format, while Java is regarded as the main programming language with open-source public reader and writer modules [61]. The data was processed into a CSV (comma-separated values) format file through selected software so that all user information including geo-locations can be listed irrespective of publication time and stored in a database as shown in Figure 2. Table 1 presents an example of a “check-in” in the CSV file. Given the heterogeneity issue, it is necessary to select only parks with more than 100 check-ins to constitute the sample of users, in order to ensure a relatively high level of representativeness.

**TABLE 1.** Example of Weibo Check-ins.

building_id	user_id	month	date	day	time	year	gender	lon	lat	Address
B2094554D06 4ABF4429	####*	08	24	Wed	0:00:03	2016	F	121.6650	31.14363	ZhaoHang_Park
B2094757D06 FA0F8409D	####*	11	11	Fri	1:24:45	2016	F	121.5244	31.265614	Jiangpu_Park
B2094757D06 FA6FB409B	####*	05	29	Mon	3:02:19	2017	M	121.3787	31.34228	Gucun_Park

**TABLE 2.** Final multiple linear regression model interpretations.

	Min	1Q	Median	3Q	Max
	-15.489	-2.308	0.325	2.184	183.555
Coefficients:	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	0.545664	0.389241	1.402	0.161213	
Baoshan	0.802558	0.438164	1.832	0.067252	.
Changning	2.739148	0.283543	9.66	< 2e-16	***
Huangpu	1.847305	0.260773	7.084	2.38E-12	***
Jingan	1.384735	0.340477	4.067	5.07E-05	***
Yangpu	1.666539	0.145104	11.485	< 2e-16	***
Mon	-0.17896	0.078598	-2.277	0.022966	*
Tue	-0.185066	0.06221	-2.975	0.00299	**
Wed	-0.182712	0.094293	-1.938	0.052891	.
Fri	-0.083539	0.063926	-1.307	0.191526	
Sat	0.003532	0.059012	0.06	0.952281	
July	0.871949	0.22472	3.88	0.00011	***
Aug	0.837961	0.202606	4.136	3.78E-05	***
Feb	1.866674	0.177615	10.51	< 2e-16	***
Mar	0.612028	0.166562	3.674	0.000249	***
Apr	0.940388	0.126142	7.455	1.71E-13	***
May	0.35173	0.140485	2.504	0.012422	*
Jun	0.829923	0.137177	6.05	1.93E-09	***
July 2014 - June 2015	0.061276	0.077947	0.786	0.431952	
July 2015 - June 2016	0.320808	0.082899	3.87	0.000115	***

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

2) STATISTICAL ANALYSIS

In order to discover the significance of explanatory variables, it was imperative to explore the predictors (explanatory variables), and the impact on the response variable (Number of Check-ins) statistically. To implement this model, we used the following regression equation:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \dots + \beta_kx_k + \epsilon. \quad (1)$$

Table 2 shows the parameters and explanatory variables used in our regression model.

$$Y = \beta_0 + \beta_1Baoshan + \beta_2Changning + \beta_3Huangpu + \beta_4Jingan + \beta_5Yangpu + \beta_6Mon + \beta_7Tue + \beta_8Wed$$

$$+ \beta_9Fri + \beta_{10}Sat + \beta_{11}July + \beta_{12}Aug + \beta_{13}Feb + \beta_{14}Mar + \beta_{15}Apr + \beta_{16}May + \beta_{17}Jun + \beta_{18}July\ 2014 - June2015 + \beta_{19}July2015 - June\ 2016 + \epsilon. \quad (2)$$

After applying a linear regression model our fitted value equation becomes

$$\hat{y} = b_0 + b_1Baoshan + b_2Changning + b_3Huangpu + b_4Jingan + b_5Yangpu + b_6Mon$$

TABLE 3. Regression statistics summary.

Residual Standard Error	Degrees of Freedom	Multiple R-Squared	Adjusted R-Squared	F-Statistic	p-value
6.512	1205	0.4873	0.4792	60.28	< 2.2e-16

TABLE 4. Analysis of variance (ANOVA).

Response: Number of Check-ins						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Baoshan	1	379	378.7	8.93	0.002861	**
Changning	1	14815	14814.7	349.3	< 2.2e-16	***
huangpu	1	8031	8030.6	189.4	< 2.2e-16	***
Jingan	1	2060	2059.6	48.56	5.25E-12	***
Yangpu	1	10885	10884.8	256.7	< 2.2e-16	***
Mon	1	45	44.8	1.057	0.3042	
Tue	1	216	215.7	5.087	0.024288	*
Wed	1	95	94.8	2.235	0.135139	
Fri	1	77	76.9	1.814	0.178312	
Sat	1	1	1.5	0.034	0.853309	
July	1	639	638.5	15.06	0.00011	***
Aug	1	814	814.4	19.2	1.28E-05	***
Feb	1	4674	4674	110.2	< 2.2e-16	***
Mar	1	339	338.6	7.985	0.004795	**
Apr	1	2959	2958.8	69.77	< 2.2e-16	***
May	1	143	142.5	3.361	0.067015	.
Jun	1	1162	1162.4	27.41	1.94E-07	***
July 2014 - June 2015	1	138	138.2	3.258	0.071333	.
July 2015 - June 2016	1	1106	1105.9	26.08	3.81E-07	***
Residuals	1205	51105	42.4			

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

$$\begin{aligned}
 &+ b_7\text{Tue} + b_8\text{Wed} + b_9\text{Fri} \\
 &+ b_{10}\text{Sat} + b_{11}\text{July} \\
 &+ b_{12}\text{Aug} + b_{13}\text{Feb} + b_{14}\text{Mar} \\
 &+ b_{15}\text{Apr} + b_{16}\text{May} + b_{17}\text{Jun} \\
 &+ b_{18}\text{July2014} - \text{June 2015} + b_{19}\text{July 2015} \\
 &- \text{June2016} + \epsilon.
 \end{aligned}
 \tag{3}$$

We applied the linear regression model using the variables that have a moderate correlation with the dependent variable, in our case Number of Check-ins is the dependent variable, having correlation values between 0.10–0.50. With regards to inference for the model, the p-value of the model’s F-statistic indicates that the model as a whole is significant. It should be noted that not all predictors have a significant p-value, as the model was developed using the highest adjusted R2 presented in Table 3. Table 2 interprets the model coefficients, in which Changning shows that, for each unit increase in the value, the Number of Check-ins is increased on average by

approximately 2.73%, with a very low p-value. Similarly, for each unit increase in the value of Huangpu, Jingan, and Yangpu, the Check-in Time is increased on average by approximately 1.84%, 1.38%, and 1.66% respectively, with a very low p-value.

Here, we can see that all independent variables are significant predictors based on their P-values [62], as shown in Table 4. For statistical analysis, we used the statistical programming language R [63] and used RStudio [64] to perform basic descriptive and regression analysis.

### 3) TEMPORAL ANALYSIS

In order to trace the changes in visitor behavior, time-stamps with check-ins can be divided into different temporal categories. We aggregated the distribution in three different year intervals, including July 2014–June 2015, July 2015–June 2016 and July 2016–June 2017. The weekly trend determines whether green parks are more beneficial

on weekends than during the week, on the basis of check-ins recorded. Similarly, the monthly distribution of check-ins between genders provides the overall trend showing which gender is more likely to use green park facilities. We used binary logistic regression to estimate the correlation and effect of the variables. In addition, we used Tableau 2019.2 for visualization techniques for exploring and analyzing relational databases and data cubes.

#### 4) SPATIAL ANALYSIS

We used KDE to construct a smooth surface density in the geographic area for check-ins hot spots. KDE is a non-parametric technique for estimating the density of a random sample of data [65]. Each data point is smoothed by KDE into small density bumps, and then all these small bumps are combined to make a final density estimate. The method of KDE has been widely used for spatial distribution [66]–[68]. KDE defines the spatial density distribution integrated with the impact of distance-decay and projects the hotspots by converting scatter point data into a continuous surface of density [69].

KDE is an evolving spatiotemporal means that has earlier been used [70], [71] to examine several characteristics of the social media (but not limited to LBSN) data analytics such as users’ online activity and movement patterns [72], check-in behavior [73], city boundary definitions [74], [75] and point-of-interest recommendations [76]. It also examines the diffusion of destinations in neighborhoods, enabling researchers to see where densely distributed destinations are and where these are more dispersed. Lastly, it attempts to create a smooth density surface in the geographic space of spatial point events [61].

KDE is an efficient measurement of the spatial structure of visitor density within a study area. KDE is a statistical approach used to estimate a smooth and continuous distribution from a limited set of observed points [77]. The data considered in our study is in the form of geo-tagged check-ins. Let “ $E$ ” be a collection of historical data for check-in, i.e.,  $E = \{e^1, \dots, e^n\}$  where  $e^i = \langle x, y \rangle$  is a check-in geo-location  $1 < i < n$ , of individual  $i$  and on time “ $t$ ” where “ $E$ ” represents the dataset we used. The sum of the functions of the kernel is scaled to create a smooth curve that is an area of the unit. This results in the following form of a KDE bivariate;

$$f_{KD}(e | E, h) = \frac{1}{n} \sum_{i=1}^n K_h(e, e^i) \tag{4}$$

$$K_h = \frac{1}{2\pi h} \exp\left(-\frac{1}{2} \sum_h (e, e^i)^t - 1\right) \tag{5}$$

where “ $e$ ” refers to the check-in location in dataset “ $E$ ” along with bandwidth “ $h$ ”. “ $h$ ” is assumed to be dependent on the estimated density “ $f_{KD}$ ”, generating a smooth density surface around “ $E$ ” at the data point “ $e^i$ ”.

ArcGIS 10.0 is used for spatial analysis to analyze the spatial distribution of check-ins in space. ArcGIS 10.0

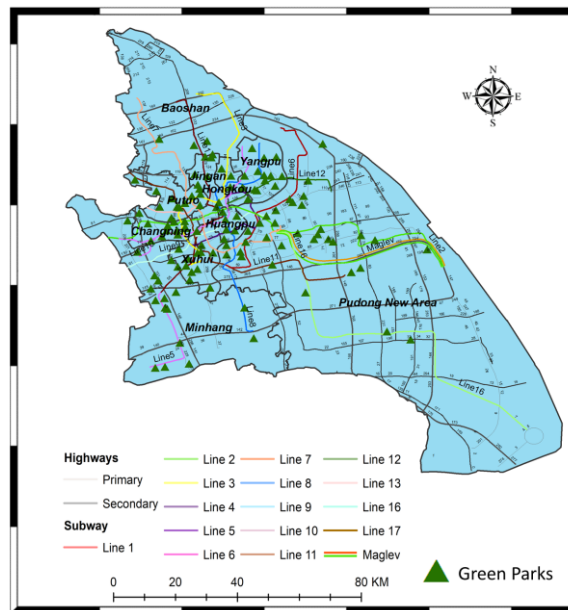


FIGURE 4. Park distribution in Shanghai.

(Environmental Systems Research Institute, Inc., Redlands, CA, USA) software with a Shanghai map generated in 2016 having Geodetic Coordinate System WGS\_1984 was used. Base-map hierarchical data also included the major transport lines (lines layer) and the administrative units (polygon layer, along with district layers), with the latest subway lines and entrances (point layer) from OpenStreetMap.

### III. RESULTS AND DISCUSSIONS

After filtering the data collected from Weibo, 115 green parks were selected for this research, as shown in Figure 4.

We used the Weibo geo-location-based check-in dataset for spatial analysis by utilizing KDE to analyze the spatial distribution of check-ins and visualizing this dataset in ArcGIS. As can be seen in Figure 5, areas shaded in red show a higher density of individuals, a higher frequency of activity and a higher concentration of social media usage. It is no wonder that the city center has large activity clusters and demonstrates an overall density of check-ins over the study period in Shanghai.

Moreover, for further analysis, we compared 3 years of check-in distribution and it can be clearly shown that there is a huge difference between the density each year. From Figure 6, it can be seen that the frequency of Weibo usage while visiting green parks has increased every year but the increment was almost doubled for the year of July 2016–June 2017, which provides the indication that public is more interested in visiting green parks for different activities.

The number of check-ins increased slightly between July 2015–June 2016 but the number of check-ins between July 2016–June 2017 has increased enormously as shown in Figure 7. The districts of Hongkou and Jingan have the lowest number of check-ins but with the passage

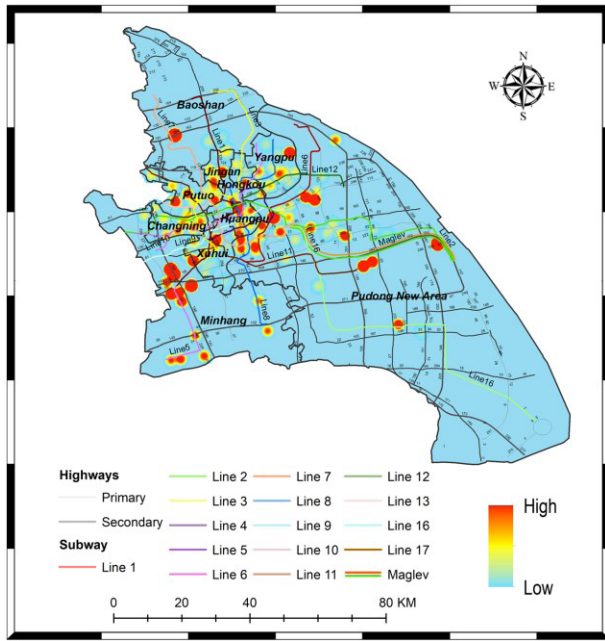


FIGURE 5. Check-in distribution for 3 years.

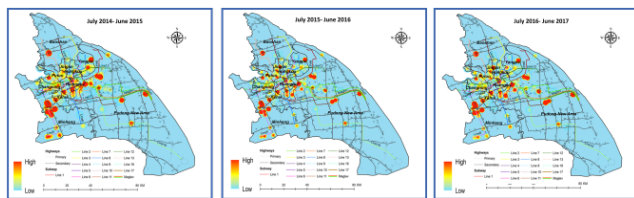


FIGURE 6. Check-in distribution per year.

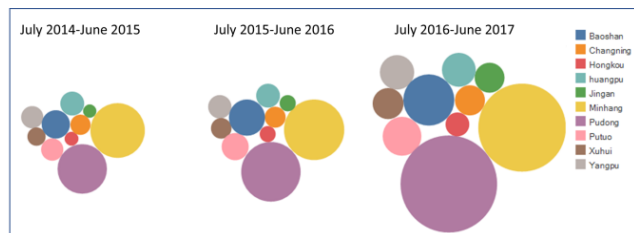


FIGURE 7. Comparison of check-in distribution in districts.

of time, the check-ins increased slightly, but for the district of Jingan, where the increase in check-ins has increased enormously. Female users are more likely to use LBSN while they are performing activities in green parks. Table 5 presents the distribution of check-ins in regard to gender difference, for each month during the 3 years of duration.

Figure 8 represents the difference between weekday and weekend check-in distribution for users in 10 districts of Shanghai. It is essential to remember that the ratio between weekdays and weekends is 5:2. The figure depicts that even though the ratio for weekends is less, the distribution of check-ins is high.

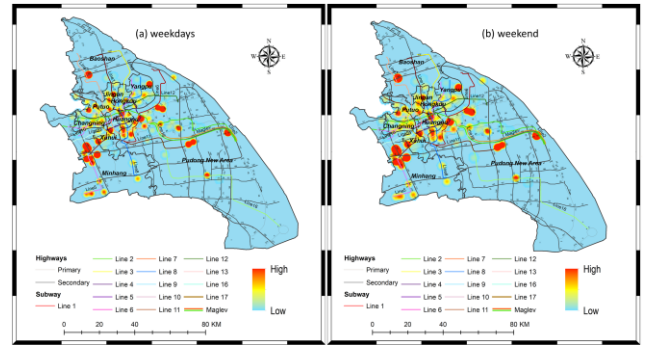


FIGURE 8. Check-in distribution for weekends and weekdays.

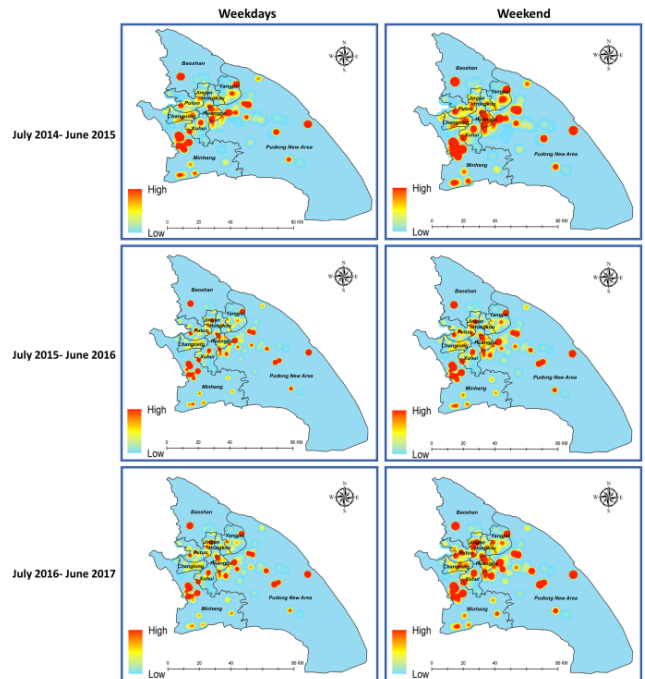


FIGURE 9. Yearly check-in distribution for weekends and weekdays.

Figure 9 shows the 3-year comparison of weekday and weekends check-in distribution of users to show behavioral trends. The decrease between check-ins on weekdays and weekends can be seen in Table 6.

The check-in distribution in Shanghai districts was calculated using the KDE method, and the density maps are visualized using ArcGIS, in both space and time. Figure 10 reveals the spatiotemporal dynamics of 10 districts of Shanghai, and it reveals the difference of density between three years. In addition to a higher density of check-ins, most of the check-ins are made in parks near to the borders of the districts next to downtown or the city-center of Shanghai.

Figure 11 represents the difference between male and female check-in distribution during the study period in 10 districts of Shanghai. There is an increasing trend in check-in density in each year. The figure reveals the ground truth that females contribute more than males in the usage of Weibo



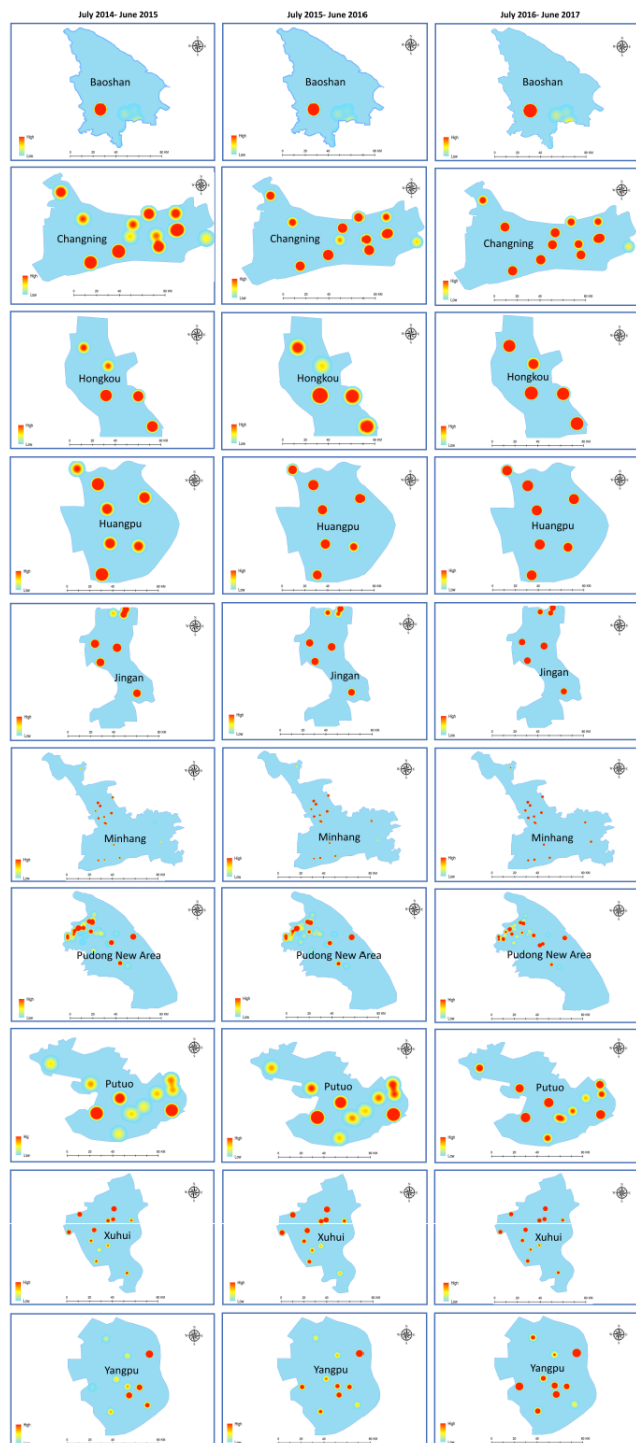


FIGURE 10. Check-in distribution in districts.

when utilizing green park facilities while engaging in different activities.

Figure 12 illustrates the gender-related check-in distribution in districts of Shanghai considered for the study. The female users' contribution in the Pudong New Area district is the highest, followed by the Minhang district, similarly, the contribution of the male users in the district of Pudong

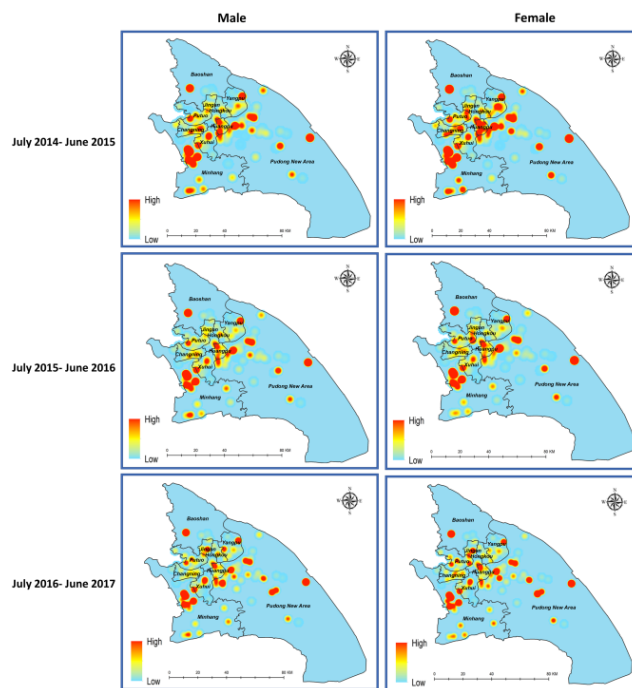


FIGURE 11. Check-in distribution by gender.

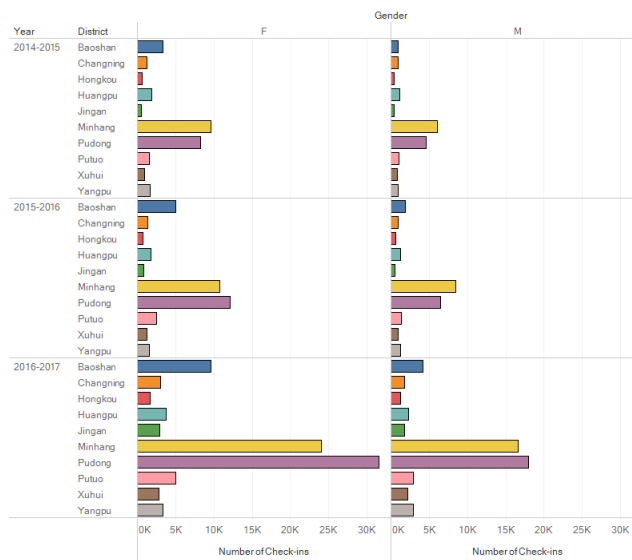


FIGURE 12. Gender-related check-in distribution in districts.

New Area is higher when compared to the other districts. In the same manner, the number of check-ins in almost every district increased year by year.

Figure 13 represents the weekly pattern for 3 years of check-ins in the dataset. The difference in the check-in behavior of male and female users during Friday, Saturday, and Sunday can be seen and is mainly due to a change in the check-in behavior of male users, which is far less than that of female users. Similarly, the same pattern can be observed from Monday to Thursday.

TABLE 5. Check-in distribution yearly, with respect to gender.

	2014-2015		2015-2016		2016-2017	
	F	M	F	M	F	M
July	0.85%	0.52%	0.92%	0.63%	2.43%	1.62%
Aug	0.85%	0.48%	1.03%	0.64%	2.39%	1.40%
Sep	0.70%	0.43%	0.86%	0.57%	2.10%	1.39%
Oct	0.98%	0.66%	1.32%	0.78%	3.15%	1.77%
Nov	0.76%	0.49%	1.29%	0.85%	2.32%	1.42%
Dec	0.72%	0.47%	1.29%	0.76%	2.45%	1.45%
Jan	0.89%	0.60%	1.32%	0.89%	1.84%	1.24%
Feb	0.69%	0.51%	1.16%	0.69%	1.89%	1.05%
Mar	0.99%	0.63%	2.19%	1.20%	2.42%	1.55%
Apr	1.57%	0.93%	1.58%	1.02%	5.28%	3.53%
May	1.70%	0.76%	1.43%	1.01%	3.32%	1.95%
Jun	1.13%	0.42%	0.76%	0.47%	5.52%	3.10%

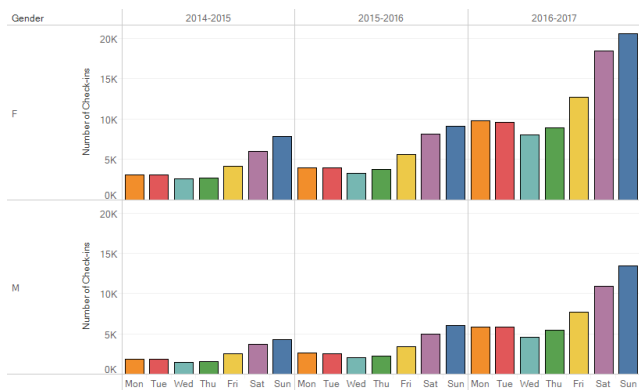


FIGURE 13. Weekly pattern for 3 years of check-ins.

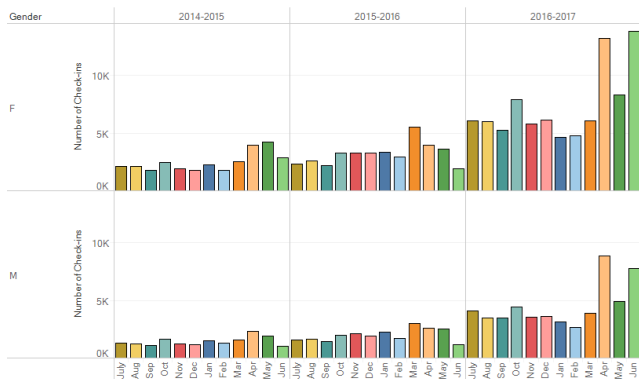


FIGURE 14. Gender check-in behavior over the months.

However, for an in-depth analysis, we observed the gender check-in behavior over the month. Figure 14 illustrates a detailed analysis of the entire months in a year, and we observe a significant increase in gender-based check-ins during March, April, and May. Moreover, it can be observed

that, during these months, the frequency was high in both genders and for July 2017, the increase in number is dramatic. The reason behind this is that these months are commonly represented as the summer season.

Table 6 illustrates the detailed distribution of check-ins with respect to the districts and days in a week. The check-in distribution districts during Friday, Saturday and Sunday is relatively high as compared to the rest of the week while the districts of Baoshan, Minhang and Pudong New Area were prominent for having a high percentage of check-ins.

Figure 15 shows the district and days trend for the overall check-ins within the study area. It can be observed that the overall pattern of check-ins is very high on weekends in Pudong New Area and Minhang district for every year, while similar patterns can be observed for weekdays as well.

The advantages of urban green parks linked with specific mechanisms for medical advantages and measures or indicators for the accessibility, availability and use of green spaces used in earlier studies [39] are addressed from the point of view of their relevance to public health and their applicability to monitor progress towards the objectives set out in international commitments, including the Parma Declaration in the European Region of the WHO and the Global Sustainable Development Goals.

First, this section discusses results focused on existing literature and scientific arguments to make it feasible for them to be meaningful and rational. The key findings are also identified by a comparison of 3 years of dataset results. This article used geotagged social media check-in data as a proxy to measuring the number of check-ins for urban green parks, which includes 10 districts of Shanghai as a case study. In comparison to the time and labor-intensive survey data, this approach is more time-efficient and also provides strong spatial coverage.

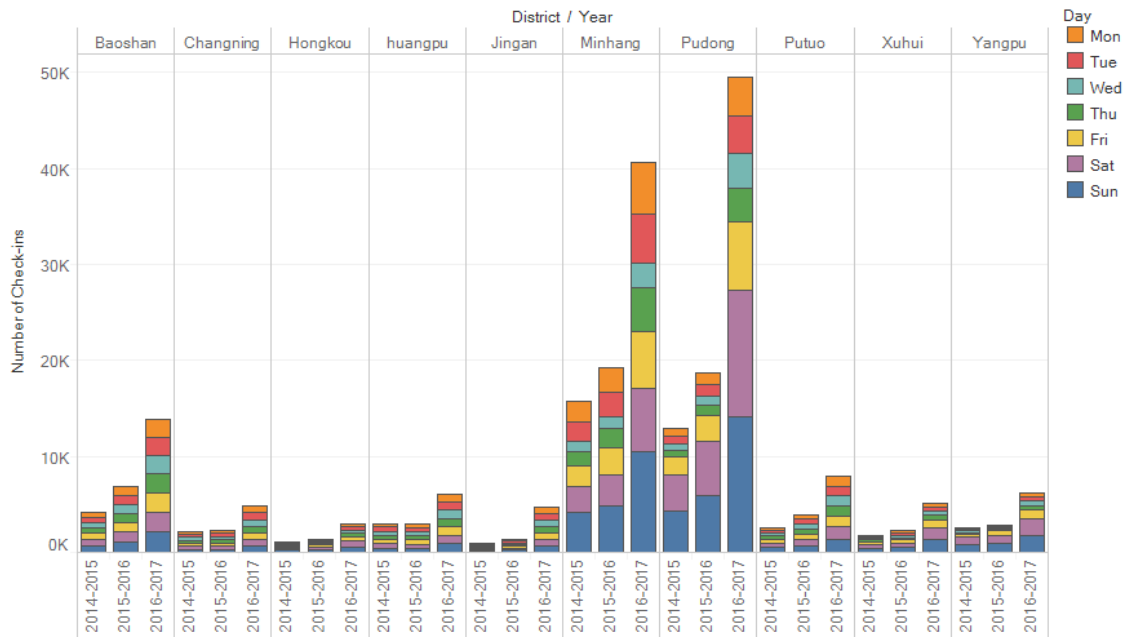


FIGURE 15. District and days trend for the overall check-in distribution.

Indeed, this approach also has its own limitations. As we do not have access to the observed number of visitation rates, we are still not able to test if there was a good correlation between both the check-in data and the observed visitation rates. The relation amongst both the Weibo check-in rate and the actual visitation may vary across green parks. The efficiency and effectiveness of the Weibo check-in data will be additionally validated when visitors' statistics are made public by the park authorities in the coming years.

Finally, based on the results, we consider the LBSN dataset to be a novel source of big data with the capability, across a new viewpoint, to be an add-on to the observation of gender-based check-in density in space and time. KDE information can make it easier to study the dynamic evolution of check-ins across both time and space. Also, the KDE results confirm that the behavior of the check-in distribution varies in fine temporal and spatial scales. These results also reveal that check-in data can represent more sophisticated phenomena and outcomes than traditional data with fine time and spatial granularity. Despite the difference of methodologies being used with different types of LBSN and datasets, both early studies [78], [79] and the current study based on gender-based check-in behavior on LBSN, draw a similar conclusion that the female users are more likely to use LBSN than male users. Although the analysis of social media data has limits, it could be seen that an exhaustive temporal, spatial and content analysis may provide important information through the grasping of general trends. Thus, this analysis serves as input for more in-depth analysis and research in the field, providing more specific objectives for decision-makers and urban planners.

The results achieved in this research may have two factors i.e., either the WHO health-related awareness report has an

impact on the substantial increase of public towards the use of green parks or the public is becoming more interested in using location-based social networks. The results presented in this research are discussed in detail as compared to earlier studies, Wendel et al., 2012 [5] analyzed the LBSN's photographs for recreational parks visitation while in our study we used LBSN's check-in data, which provides a more efficient and accurate way to carry out the research. Furthermore, this study investigated spatial variations in urban green parks, which was presented by Wenping et al., 2017 [80], and, in addition, the temporal variations, along-with gender differences, are also provided. The study has been carried-out on a large-scale dataset to present more in-depth details on the influence of public behavior towards parks for healthy activities.

There are also some negative effects of green parks, for example, the evidence of an association among urban greenery, asthma and allergies are rather inconclusive. Lovasi et al., 2008 [81] observed that children living in an area with more gardens and trees in New York City had a low prevalence of asthma. A subsequent cohort study involving minor children in New York City did not demonstrate a hypothesized protective impact and in reality, showed a positive relationship between tree cover and infectious sensitization to tree pollen and the disease of asthma in children [58]. Another study in the United States revealed that pollen is connected with urban green parks and trees were one of the self-reported factors that cause asthma in Philadelphia [82]. Living near to green areas can be associated with high exposure to herbicides and pesticides, particularly when used inappropriately and at extreme levels. Green space health risks include arthropod-borne illnesses for example ticks (e.g., tick-borne encephalitis and Lyme

**TABLE 6.** Representation of check-ins with respect to districts and days in a week.

Year	Day	Baoshan	Changning	Hongkou	Huangpu	Jingan	Minhang	Pudong New Area	Putuo	Xuhui	Yangpu
July 2014 - June 2015	Sun	0.28%	0.13%	0.09%	0.18%	0.06%	1.68%	1.73%	0.20%	0.17%	0.33%
	Mon	0.22%	0.12%	0.04%	0.16%	0.06%	0.85%	0.32%	0.13%	0.05%	0.07%
	Tue	0.23%	0.12%	0.04%	0.17%	0.06%	0.79%	0.32%	0.12%	0.07%	0.06%
	Wed	0.23%	0.13%	0.05%	0.18%	0.05%	0.45%	0.28%	0.12%	0.07%	0.08%
	Thu	0.22%	0.12%	0.05%	0.17%	0.05%	0.60%	0.26%	0.12%	0.07%	0.06%
	Fri	0.24%	0.14%	0.06%	0.17%	0.05%	0.85%	0.74%	0.15%	0.12%	0.14%
	Sat	0.28%	0.12%	0.08%	0.18%	0.05%	1.04%	1.50%	0.20%	0.17%	0.30%
July 2015 - June 2016	Sun	0.43%	0.13%	0.12%	0.17%	0.08%	1.95%	2.35%	0.27%	0.19%	0.37%
	Mon	0.39%	0.14%	0.05%	0.17%	0.08%	1.02%	0.45%	0.18%	0.09%	0.06%
	Tue	0.38%	0.14%	0.05%	0.16%	0.08%	0.98%	0.47%	0.19%	0.10%	0.06%
	Wed	0.37%	0.13%	0.06%	0.18%	0.08%	0.51%	0.41%	0.21%	0.10%	0.07%
	Thu	0.39%	0.12%	0.06%	0.17%	0.07%	0.82%	0.41%	0.20%	0.08%	0.07%
	Fri	0.39%	0.13%	0.09%	0.18%	0.08%	1.11%	1.11%	0.23%	0.13%	0.17%
	Sat	0.42%	0.14%	0.11%	0.17%	0.08%	1.28%	2.25%	0.27%	0.20%	0.34%
July 2016 - June 2017	Sun	0.83%	0.27%	0.22%	0.35%	0.27%	4.21%	5.63%	0.53%	0.54%	0.70%
	Mon	0.75%	0.26%	0.13%	0.33%	0.27%	2.14%	1.57%	0.42%	0.16%	0.18%
	Tue	0.74%	0.29%	0.13%	0.33%	0.27%	2.07%	1.58%	0.40%	0.16%	0.18%
	Wed	0.77%	0.27%	0.13%	0.35%	0.27%	1.02%	1.43%	0.43%	0.18%	0.19%
	Thu	0.77%	0.28%	0.14%	0.33%	0.28%	1.81%	1.41%	0.40%	0.18%	0.17%
	Fri	0.80%	0.27%	0.20%	0.35%	0.26%	2.33%	2.81%	0.44%	0.30%	0.37%
	Sat	0.85%	0.27%	0.25%	0.37%	0.27%	2.63%	5.29%	0.55%	0.51%	0.71%

disease), mosquitoes (e.g., Chikungunya fever and Dengue fever).

To the best of our knowledge, no previous research has been conducted for a comparison of a large-scale dataset for observing the behavior of green parks visitors, which can provide awareness about the minimization of serious diseases in comparison to WHO report [83] published in 2016 by performing healthy activities in green parks. This research might be the first case study using Weibo data to analyze the use of urban green parks concerning check-in behavior for a large number of green parks in Shanghai. The broad spatial coverage of this study provides useful information that could enhance urban green space planning and development in other cities. There is significant potential for encouraging moderate to vigorous physical activity among urban residents with the effective use of urban green areas. These results emphasize the significance of the use of urban green parks to improve the health of residents living in cities around the world.

#### IV. CONCLUSION

Urban green parks are areas where people can participate in health-improving physical activities. Therefore, studies

exploring the relationship between green parks and their use are important for the restoration of human health. We investigated spatial, temporal and affective patterns of 3 years of visits to public parks using Weibo data for health purposes from frequent users in Shanghai and compared the check-ins of each year based upon weekly and monthly time frames to observe how the behavior of visitors is changing each year. Our findings show that the number of visitors is increasing year by year, and this LBSN based study draws a conclusion that female visitors are more likely to visit green parks than male visitors. These results emphasize the significance of concentrating on the use of urban green spaces to improve the health of urban residents around the world. The main objectives of this research were to raise awareness of green parks association with health, to provide evidence for policymakers on the advantages of providing green space access to urban residents and to provide the urban cities with systematic new approaches for quantifying and monitoring their access to green space to enhance the healthy activities among citizens.

#### CONFLICTS OF INTEREST

The authors declare no conflict of interests.

## REFERENCES

- [1] K. L. Hand, C. Freeman, P. J. Seddon, M. R. Recio, A. Stein, and Y. Van Heezik, "The importance of urban gardens in supporting children's biophilia," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 2, pp. 274–279, Jan. 2017.
- [2] H. Ullah, W. Wan, S. A. Haidery, N. U. Khan, Z. Ebrahimpour, and T. Luo, "Analyzing the spatiotemporal patterns in green spaces for urban studies using location-based social media data," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 11, p. 506, Nov. 2019.
- [3] T. C. Daniel, A. Muhar, A. Arnberger, O. Aznar, J. W. Boyd, K. M. Chan, R. Costanza, T. Elmqvist, C. G. Flint, and P. H. Gobster, "Contributions of cultural services to the ecosystem services agenda," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 23, pp. 8812–8819, 2012.
- [4] D. A. Cohen, T. Marsh, S. Williamson, K. P. Derose, H. Martinez, C. Setodji, and T. L. McKenzie, "Parks and physical activity: Why are some parks used more than others?" *Preventive Med.*, vol. 50, pp. S9–S12, Jan. 2010.
- [5] H. E. W. Wendel, R. K. Zarger, and J. R. Mihelcic, "Accessibility and usability: Green space preferences, perceptions, and barriers in a rapidly urbanizing city in Latin America," *Landscape Urban Planning*, vol. 107, no. 3, pp. 272–282, Sep. 2012.
- [6] C. Sessions, S. A. Wood, S. Rabotyagov, and D. M. Fisher, "Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs," *J. Environ. Manage.*, vol. 183, pp. 703–711, Dec. 2016.
- [7] H. Tenkanen, E. Di Minin, V. Heikinheimo, A. Hausmann, M. Herbst, L. Kajala, and T. Toivonen, "Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 17615.
- [8] M. Cetin, "Determining the bioclimatic comfort in Kastamonu City," *Environ. Monit. Assessment*, vol. 187, p. 640, Oct. 2015.
- [9] A. C. K. Lee and R. Maheswaran, "The health benefits of urban green spaces: A review of the evidence," *J. Public Health*, vol. 33, no. 2, pp. 212–222, Jun. 2011.
- [10] S. M. A. Haq, "Urban green spaces and an integrative approach to sustainable environment" *J. Environ. Protection*, vol. 2, no. 5, p. 601, 2011.
- [11] E. Amoly, P. Dadvand, J. Fornis, M. L. López-Vicente, X. Basagaña, J. Julvez, M. Alvarez-Pedrerol, M. J. Nieuwenhuijsen, and J. Sunyer, "Green and blue spaces and behavioral development in Barcelona schoolchildren: The BREATHE project," *Environ. Health Perspect.*, vol. 122, no. 12, pp. 1351–1358, Dec. 2014.
- [12] B. Giles-Corti, M. H. Broomhall, M. Knuiaman, C. Collins, K. Douglas, K. Ng, A. Lange, and R. J. Donovan, "Increasing walking: How important is distance to, attractiveness, and size of public open space?" *Amer. J. Preventive Med.*, vol. 28, no. 2, pp. 169–176, 2005.
- [13] A. Chiesura, "The role of urban parks for the sustainable city," *Landscape Urban Planning*, vol. 68, no. 1, pp. 129–138, May 2004.
- [14] *Transforming Our World: The 2030 Agenda for Sustainable Development*. Accessed: Jun. 14, 2019. [Online]. Available: <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>
- [15] H. Roberts, J. Sadler, and L. Chapman, "Using Twitter to investigate seasonal variation in physical activity in urban green space," *Geo, Geography Environ.*, vol. 4, no. 2, Jul. 2017, Art. no. e00041.
- [16] R. A. Kolcsár and P. Szilassi, "Assessing accessibility of urban green spaces based on isochrone maps and street resolution population data through the example of Zalaegerszeg, Hungary," *Carpathian J. Earth Environ. Sci.*, vol. 13, no. 1, pp. 31–36, 2018.
- [17] J. Maas, R. A. Verheij, S. De Vries, P. Spreeuwenberg, F. G. Schellevis, and P. P. Groenewegen, "Morbidity is related to a green living environment," *J. Epidemiol. Community Health*, vol. 63, no. 12, pp. 967–973, Dec. 2009.
- [18] A. E. Bauman and F. C. Bull, *Environmental Correlates of Physical Activity and Walking in Adults and Children: A Review of Reviews*. London, U.K.: National Institute of Health and Clinical Excellence, Feb. 2007.
- [19] E. D. Ekkel and S. De Vries, "Nearby green space and human health: Evaluating accessibility metrics," *Landscape Urban Planning*, vol. 157, pp. 214–220, Jan. 2017.
- [20] G. R. McCormack, M. Rock, A. M. Toohy, and D. Hignell, "Characteristics of urban parks associated with park use and physical activity: A review of qualitative research," *Health Place*, vol. 16, no. 4, pp. 712–726, Jul. 2010.
- [21] A. T. Kaczynski, J. Schipperijn, J. A. Hipp, G. M. Besenyi, S. A. W. Stanis, S. M. Hughey, and S. Wilcox, "ParkIndex: Development of a standardized metric of park access for research and planning," *Preventive Med.*, vol. 87, pp. 110–114, Jun. 2016.
- [22] F. Erkip, "The distribution of urban public services: The case of parks and recreational services in Ankara," *Cities*, vol. 14, no. 6, pp. 353–361, Dec. 1997.
- [23] S. Macintyre, L. Macdonald, and A. Ellaway, "Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland," *Social Sci. Med.*, vol. 67, no. 6, pp. 900–914, 2008.
- [24] H.-S. Lee, "Examining neighborhood influences on leisure-time walking in older Korean adults using an extended theory of planned behavior," *Landscape Urban Planning*, vol. 148, pp. 51–60, Apr. 2016.
- [25] J. Byrne, J. Wolch, and J. Zhang, "Planning for environmental justice in an urban national park," *J. Environ. Planning Manage.*, vol. 52, no. 3, pp. 365–392, Apr. 2009.
- [26] G. Lindsey, M. Maraj, and S. Kuan, "Access, equity, and urban greenways: An exploratory investigation," *Prof. Geographer*, vol. 53, no. 3, pp. 332–346, Aug. 2001.
- [27] D. Wang, G. Brown, and I. Mateo-Babiano, "Beyond proximity: An integrated model of accessibility for public parks," *Asian J. Social Sci. Hum.*, vol. 2, no. 3, pp. 486–498, 2013.
- [28] D. Wang, G. Brown, and Y. Liu, "The physical and non-physical factors that influence perceived access to urban parks," *Landscape Urban Planning*, vol. 133, pp. 53–66, Jan. 2015.
- [29] H. A. Carlos, X. Shi, J. Sargent, S. Tanski, and E. M. Berke, "Density estimation and adaptive bandwidths: A primer for public health practitioners," *Int. J. Health Geographics*, vol. 9, no. 1, p. 39, 2010.
- [30] C. L. Szwarcwald, C. L. T. D. Andrade, and F. I. Bastos, "Income inequality, residential poverty clustering and infant mortality: A study in Rio de Janeiro, Brazil," *Social Sci. Med.*, vol. 55, no. 12, pp. 2083–2092, Dec. 2002.
- [31] D. A. Moore and T. E. Carpenter, "Spatial analytical methods and geographic information systems: Use in health research and epidemiology," *Epidemiol. Rev.*, vol. 21, no. 2, pp. 143–161, Jan. 1999.
- [32] F. Li, F. Zhang, X. Li, P. Wang, J. Liang, Y. Mei, W. Cheng, and Y. Qian, "Spatiotemporal patterns of the use of urban green spaces and external factors contributing to their use in central Beijing," *Int. J. Environ. Res. Public Health*, vol. 14, no. 3, p. 237, Feb. 2017.
- [33] K. Glanz, "Measuring food environments: A historical perspective," *Amer. J. Preventive Med.*, vol. 36, no. 4, pp. S93–S98, 2009.
- [34] J. Feng, T. A. Glass, F. C. Curriero, W. F. Stewart, and B. S. Schwartz, "The built environment and obesity: A systematic review of the epidemiologic evidence," *Health Place*, vol. 16, no. 2, pp. 175–190, Mar. 2010.
- [35] F. Li, P. Harmer, B. J. Cardinal, M. Bosworth, and D. Johnson-Shelton, "Obesity and the built environment: Does the density of neighborhood fast-food outlets matter?" *Amer. J. Health Promotion*, vol. 23, no. 3, pp. 203–209, 2009.
- [36] D. A. Crawford, A. F. Timperio, J. A. Salmon, L. Baur, B. Giles-Corti, R. J. Roberts, M. L. Jackson, N. Andrianopoulos, and K. Ball, "Neighbourhood fast food outlets and obesity in children and adults: The CLAN Study," *Int. J. Pediatric Obesity*, vol. 3, no. 4, pp. 249–256, Jan. 2008.
- [37] L. M. Seliske, W. Pickett, W. F. Boyce, and I. Janssen, "Association between the food retail environment surrounding schools and overweight in Canadian youth," *Public Health Nutr.*, vol. 12, no. 9, pp. 1384–1391, Sep. 2009.
- [38] L. V. Moore, A. V. D. Roux, and S. Brines, "Comparing perception-based and geographic information system (GIS)-based characterizations of the local food environment," *J. Urban Health*, vol. 85, no. 2, pp. 206–216, Mar. 2008.
- [39] *Global Recommendations on Physical Activity for Health*, World Health Org., Geneva, Switzerland, 2010.
- [40] *Physical Activity Promotion in Socially Disadvantaged Groups: Principles for Action*, PHAN Work Package 4 Final Rep., WHO Regional Office Eur. Copenhagen, World Health Org., Geneva, Switzerland, 2013.
- [41] T. Hartig, R. Mitchell, S. De Vries, and H. Frumkin, "Nature and health," *Annu. Rev. Public Health*, vol. 35, pp. 207–228, Mar. 2014.
- [42] D. Krewski and D. Rainham, "Ambient air pollution and population health: Overview," *J. Toxicol. Environ. Health, A*, vol. 70, nos. 3–4, pp. 275–283, Feb. 2007.
- [43] A. Tiwary, D. Sinnett, C. Peachey, Z. Chalabi, S. Vardoulakis, T. Fletcher, G. Leonardi, C. Grundy, A. Azapagic, and T. R. Hutchings, "An integrated tool to assess the role of new planting in PM10 capture and the human health benefits: A case study in London," *Environ. Pollution*, vol. 157, no. 10, pp. 2645–2653, Oct. 2009.

- [44] J. Barton and J. Pretty, "What is the best dose of nature and green exercise for improving mental health? A multi-study analysis," *Environ. Sci. Technol.*, vol. 44, no. 10, pp. 3947–3955, May 2010.
- [45] S. De Vries, R. A. Verheij, P. P. Groenewegen, and P. Spreeuwenberg, "Natural environments—Healthy environments? An exploratory analysis of the relationship between greenspace and health," *Environ. Plan A*, vol. 35, no. 10, pp. 1717–1731, Oct. 2003.
- [46] R. Mitchell and F. Popham, "Greenspace, urbanity and health: Relationships in England," *J. Epidemiol. Community Health*, vol. 61, no. 8, pp. 681–683, Aug. 2007.
- [47] C. Paquet, T. P. Orschulok, N. T. Coffee, N. J. Howard, G. Hugo, A. W. Taylor, R. J. Adams, and M. Daniel, "Are accessibility and characteristics of public open spaces associated with a better cardiometabolic health?" *Landscape Urban Planning*, vol. 118, pp. 70–78, Oct. 2013.
- [48] D. R. Dengel, M. O. Hearst, J. H. Harmon, A. Forsyth, and L. A. Lytle, "Does the built environment relate to the metabolic syndrome in adolescents?" *Health Place*, vol. 15, no. 4, pp. 946–951, Dec. 2009.
- [49] B.-J. Park, Y. Tsunetsugu, T. Kasetani, T. Morikawa, T. Kagawa, and Y. Miyazaki, "Physiological effects of forest recreation in a young conifer forest in Hinokage Town, Japan," *Silva Fennica*, vol. 43, no. 2, pp. 291–301, 2009.
- [50] G.-X. Mao, Y.-B. Cao, X.-G. Lan, Z.-H. He, Z.-M. Chen, Y.-Z. Wang, X.-L. Hu, Y.-D. Lv, G.-F. Wang, and J. Yan, "Therapeutic effect of forest bathing on human hypertension in the elderly," *J. Cardiol.*, vol. 60, no. 6, pp. 495–502, Dec. 2012.
- [51] *Department of Economic and Social Affairs. World Urbanization Prospects the 2014 Revision*, United Nations, New York, NY, USA, 2014.
- [52] (2019). *Demographia World Urban Area*. Accessed: Jun. 14, 2019. [Online]. Available: <http://www.demographia.com/db-worldua.pdf>
- [53] Y. Xiao, Z. Wang, Z. Li, and Z. Tang, "An assessment of urban park access in Shanghai—implications for the social equity in urban China," *Landscape Urban Planning*, vol. 157, pp. 383–393, Jan. 2017.
- [54] X. Xiong, C. Jin, H. Chen, and L. Luo, "Using the fusion proximal area method and gravity method to identify areas with physician shortages," *PLoS ONE*, vol. 11, no. 10, Oct. 2016, Art. no. e0163504.
- [55] J. Shen and G. Kee, *Development and Planning in Seven Major Coastal Cities in Southern and Eastern China*. Cham, Switzerland: Springer, 2017.
- [56] *Weibo Statistics*. Accessed: Jun. 10, 2019. [Online]. Available: <http://ir.weibo.com/news-releases/news-release-details/weibo-corporation-report-fourth-quarter-and-fiscal-year-2018>
- [57] Y. Shen, F. Sun, and Y. Che, "Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai," *Urban Forestry Urban Greening*, vol. 27, pp. 59–68, Oct. 2017.
- [58] Y. Shen and K. Karimi, "Urban function connectivity: Characterisation of functional urban streets with social media check-in data," *Cities*, vol. 55, pp. 9–21, Jun. 2016.
- [59] B. L. Keeler, S. A. Wood, S. Polasky, C. Kling, C. T. Filstrup, and J. A. Downing, "Recreational demand for clean water: Evidence from geo-tagged photographs by visitors to lakes," *Frontiers Ecology Environment*, vol. 13, no. 2, pp. 76–81, Mar. 2015.
- [60] Y. Wang, T. Wang, X. Ye, J. Zhu, and J. Lee, "Using social media for emergency response and urban sustainability: A case study of the 2012 Beijing rainstorm," *Sustainability*, vol. 8, no. 1, p. 25, Dec. 2015.
- [61] Z. Xie and J. Yan, "Kernel density estimation of traffic accidents in a network space," *Comput., Environ. Urban Syst.*, vol. 32, no. 5, pp. 396–406, Sep. 2008.
- [62] S. Abidi, M. Hussain, Y. Xu, and W. Zhang, "Prediction of confusion attempting algebra homework in an intelligent tutoring system through machine learning techniques for educational sustainable development," *Sustainability*, vol. 11, no. 1, p. 105, Dec. 2018.
- [63] *R Language*. Accessed: May 12, 2019. [Online]. Available: <https://cran.r-project.org>
- [64] *R Studio*. Accessed: May 12, 2019. [Online]. Available: <https://www.rstudio.com>
- [65] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Evanston, IL, USA: Routledge, 2018.
- [66] A. R. Maroko, J. A. Maantay, N. L. Sohler, K. L. Grady, and P. S. Arno, "The complexities of measuring access to parks and physical activity sites in New York City: A quantitative and qualitative approach," *Int. J. Health Geographics*, vol. 8, no. 1, p. 34, 2009.
- [67] W. Yu, T. Ai, and S. Shao, "The analysis and delimitation of central business district using network kernel density estimation," *J. Transp. Geography*, vol. 45, pp. 32–47, May 2015.
- [68] W. Yu, T. Ai, Y. He, and S. Shao, "Spatial co-location pattern mining of facility points-of-interest improved by network neighborhood and distance decay effects," *Int. J. Geographical Inf. Sci.*, vol. 31, no. 2, pp. 280–296, Feb. 2017.
- [69] L. Ying, Z. Shen, J. Chen, R. Fang, X. Chen, and R. Jiang, "Spatiotemporal patterns of road network and road development priority in three parallel rivers region in Yunnan, China: An evaluation based on modified kernel distance estimate," *Chin. Geogr. Sci.*, vol. 24, no. 1, pp. 39–49, Feb. 2014.
- [70] C. Wu, X. Ye, F. Ren, Y. Wan, P. Ning, and Q. Du, "Spatial and social media data analytics of housing prices in Shenzhen, China," *PLoS ONE*, vol. 11, no. 10, Oct. 2016, Art. no. e0164553.
- [71] T. L. King, R. J. Bentley, L. E. Thornton, and A. M. Kavanagh, "Using kernel density estimation to understand the influence of neighbourhood destinations on BMI," *BMJ Open*, vol. 6, no. 2, Feb. 2016, Art. no. e008878.
- [72] S. Hasan, X. Zhan, and S. V. Ukkusuri, "Understanding urban human activity and mobility patterns using large-scale location-based data from online social media," in *Proc. 2nd ACM SIGKDD Int. Workshop Urban Comput. (UrbComp)*, Chicago, IL, USA, Aug. 2013, p. 6.
- [73] L. Li, M. F. Goodchild, and B. Xu, "Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr," *Cartography Geographic Inf. Sci.*, vol. 40, no. 2, pp. 61–77, Mar. 2013.
- [74] Y. Sun, H. Fan, M. Li, and A. Zipf, "Identifying the city center using human travel flows generated from location-based social networking data," *Environ. Planning B, Planning Des.*, vol. 43, no. 3, pp. 480–498, May 2016.
- [75] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong, "Discovering urban functional zones using latent activity trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 3, pp. 712–725, Mar. 2015.
- [76] H. Li, Y. Ge, R. Hong, H. Zhu, "Point-of-interest recommendations: Learning potential check-ins from friends," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 975–984.
- [77] M. Maia, J. Almeida, and V. Almeida, "Identifying user behavior in online social networks," in *Proc. 1st Workshop Social Netw. Syst.*, Glasgow, Scotland, Apr. 2008, pp. 1–6.
- [78] H. S. Hwang and E. K. Choi, "Exploring gender differences in motivations for using Sina Weibo," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 3, pp. 1–13, 2016.
- [79] M. Chun, "The affective/cognitive involvement and satisfaction according to the usage motivations of social network services," *Manage. Inf. Syst. Rev.*, vol. 31, no. 2, pp. 21–39, Jun. 2012.
- [80] W. Liu, W. Chen, and C. Dong, "Spatial decay of recreational services of urban parks: Characteristics and influencing factors," *Urban Forestry Urban Greening*, vol. 25, pp. 130–138, Jul. 2017.
- [81] G. S. Lovasi, J. W. Quinn, K. M. Neckerman, M. S. Perzanowski, and A. Rundle, "Children living in areas with more street trees have lower prevalence of asthma," *J. Epidemiol. Community Health*, vol. 62, no. 7, pp. 647–649, 2008.
- [82] S. Keddem, F. K. Barg, K. Glanz, T. Jackson, S. Green, and M. George, "Mapping the urban asthma experience: Using qualitative GIS to understand contextual factors affecting asthma control," *Social Sci. Med.*, vol. 140, pp. 9–17, Sep. 2015.
- [83] *Urban Green Spaces and Health*, World Health Org., Geneva, Switzerland, 2016.



**HIDAYAT ULLAH** received the master's degree from Preston University, Islamabad, Pakistan, in 2015. He is currently pursuing the Ph.D. degree with the School of Communication and Information Engineering, Shanghai University, Shanghai, China. He has authored three SCI indexed articles in internationally renowned journals. His research interests mainly focused on big data, social media, environmental science, smart city, and urban planning. He has been involved in several research

projects, including projects of the National Natural Science Foundation of China and the Shanghai Science and Technology Commission. He is also a member of the Institute of Smart City, Shanghai University. He has been awarded the CSC Scholarship (China Scholarship Council) to study at Shanghai University. He has the honor to Co-Chair the IEEE Conference ICALIP 2017.



**WANGGEN WAN** (Senior Member, IEEE) received the Ph.D. degree from Xidian University, China, in 1992. From 1991 to 1992, he was a Visiting Scholar at the Department of Computer Engineering, Minsk Radio Engineering Institute, USSR. He was a Postdoctoral Research Fellow at the Department of Information and Control Engineering, Xi'an Jiao-Tong University, from 1993 to 1995. He was a Visiting Scholar at the Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology, from 1998 to 1999. He was a Visiting Professor and the Section Head of the Multimedia Innovation Center, The Hong Kong Polytechnic University, from 2000 to 2004. He has been a Full Professor with the School of Communication and Information Engineering, Shanghai University, since 2004. He is also the Director of the Institute of Smart City, Shanghai University, and the Dean of International Office, Shanghai University. His research interests include computer graphics, video and image processing, and data mining. He has authored over 200 academic articles on international journals and conferences, and he has been involved in over 30 research projects as the Principal Investigator. He is the Vice-Chair of the IEEE CIS Shanghai Chapter and the Chair of the IET Shanghai Local Network. He is a Fellow of the IET, and an ACM Professional Member. He has been the Co-Chairman for many well-known international conferences, since 2008.



**SAQIB ALI HAIDERY** received the bachelor's degree in computer engineering from the Department of Electrical Engineering, COMSATS Institute of Information Technology, Islamabad, Pakistan. He is currently pursuing the master's degree with the School of Communication and Information Engineering, Shanghai University, Shanghai, China. He is also a member of the Institute of Smart City, Shanghai University. He has authored two SCI Indexed and two EI indexed articles in internationally renowned journals and conferences in the areas of applications of big data and smart city. His research interests are mainly focused on big data, social media, data science, smart city, and urban planning. He has been involved in several research projects, including projects of the General Program of National Natural Science Foundation of China and the Shanghai Science and Technology Commission. He has been awarded the CSC Scholarship (China Scholarship Council) to study at Shanghai University, from September 2015 to August 2018.



**NAIMAT ULLAH KHAN** received the master's degree from Preston University, Islamabad, Pakistan, in 2015. He is currently pursuing the Ph.D. degree with the School of Communication and Information Engineering, Shanghai University, Shanghai, China. He is also a member of the Institute of Smart City, Shanghai University. He has authored three SCI Indexed articles in internationally renowned journals and conferences. His research interests are mainly focused on big data, social media, tourism, data science, and urban planning. He has been involved in several research projects, including projects of the National Natural Science Foundation of China and the Shanghai Science and Technology Commission. He has been awarded the Shanghai Government Scholarship to study at Shanghai University.



**ZEINAB EBRAHIMPOUR** was born in Sary, Mazandaran, Iran, in 1991. She received the B.S. degree in software engineering and the M.S. degree in computer science from the Hadaf Higher Education Institute, in 2014 and 2016, respectively, with a GPA of 18.88/20. She is currently pursuing the Ph.D. degree in communication and information engineering with Shanghai University, Shanghai, China. In 2016, she joined the Institute of Smart City, Shanghai University, starting her research on data mining and big data analysis and cooperating with other research groups in the Univeisidad de las Americas Puebla, Mexico. Her research interests include big data, urban planning, human mobility, and artificial intelligence in social media data analysis. Her honors include the talented student in master's degree and the Co-Chair of the IEEE Conference ICSSC 2016 and ICALIP 2017.



**A. A. M. MUZAHID** received the B.Sc. degree in electronics and telecommunications engineering from Daffodil International University, Dhaka, Bangladesh, in 2011, and the M.E. degree in communication and information engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2016. He is currently pursuing the Ph.D. degree in communication and information systems with Shanghai University, Shanghai, China. He is a research member of the Institute of Smart City, Shanghai University, and leading the "Computer Vision and 3D Virtual Reality" Research Team. His current research interests include 3D shape analysis, computer vision, and 3D virtual reality.

...