Beyond home neighborhood
Mobility, activity and temporal variation of socio-spatial segregation

Xian, Shi; Qi, Zhixin; Yip, Ngai-ming

Published in:
Journal of Transport Geography

Published: 01/02/2022

Document Version:
Post-print, also known as Accepted Author Manuscript, Peer-reviewed or Author Final version

Published version (DOI):
10.1016/j.jtrangeo.2022.103304

Publication details:
Xian, S., Qi, Z., & Yip, N. (2022). Beyond home neighborhood: Mobility, activity and temporal variation of socio-spatial segregation. *Journal of Transport Geography*, *99*, Article 103304. https://doi.org/10.1016/j.jtrangeo.2022.103304

Beyond Home Neighborhood: Mobility, Activity and Temporal Variation of Socio-spatial Segregation

Shi Xian [a], Zhixin Qi [b, *], Ngai-ming Yip [c]

a Guangzhou University – The University of Birmingham Joint Research Centre of Service Industries and Urban Development, School of Geographical Sciences and Remote Sensing, Guangzhou University, Guangzhou, P. R. China

b Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, P. R. China

c Department of Public Policy, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong SAR, P.R. China

Funding information

This work was supported by the National Key R&D Program of China (Grant No. 2019YFB2103102), the Guangdong Planning Project of Philosophy and Social Science (Grant No.GD21CGL15), the Pearl River Talent Recruitment Program (Grant No.20170142) and the National Natural Science Foundation of China (Grant No.41701189).

Beyond Home Neighborhood: Mobility, Activity and Temporal Variation of Socio-spatial Segregation

Abstract

Recent studies on socio-spatial segregation have revealed the uneven segregation experiences of individuals within their daily life contexts. However, little is known about its temporal variations across the week. The advancement of GIS and GPS tracking technology also poses methodological challenges in processing rich mobility–activity data efficiently in identifying the socio-spatial segregation patterns. With data collected by a mobile phone app that ran on the participants' mobile phone for a whole week, this paper integrates the spatial, temporal, mobility, and activity dimensions with the demographic data to segregation patterns of the participants and assesses segregation at the individual level. Our findings indicate that the socio-spatial segregation level decreased in the daytime and increased at night, and this pattern was consistent across a week. However, no significant differences are found between different age groups, occupation, housing types and home neighborhood types. To improve the efficiency of data processing, this paper employs decision tree algorithms supplemented by the analysis of variance and Tukey's honestly significant difference test to identify meaningful mobility–activity patterns with significant intergroup differences. It is able to pinpoint temporal and spatial activity-mobility patterns that crosscut home location, location of workplace, and socioeconomic status. It also helps connect residential segregation and segregation that goes beyond the home neighborhoods.

Key words

Socio-spatial segregation, mobility, activity, housing tenure status, decision tree algorithms

1. Introduction

Socio-spatial segregation has long been the theme of numerous studies and debates since the early works by the Chicago School, including that of Park and Burgess (1925). This concept concerns the extent to which people of different socioeconomic status are segregated, reflecting lack of contact, interactions, or participation in mutual activities. Academic research and debates revolve around the negative or positive effects of socio-spatial segregation, and the discussions are usually related to social disadvantages, discrimination, and inequalities. The concept has since been extended to refer to the voluntary or involuntary condition of isolation in an urban area of a particular social group (Yuan, 1963; Lieberson and Carter, 1982). At the early stage, the ethnic identity of individuals was the main focus of most socio-spatial segregation studies that conducted extensive empirical analyses in the United States. Thereafter, more studies were conducted in Europe and other areas, with the major concerns centered on social economic status, including the composition of social classes, gender, housing quality, and age (Van Kempen, 1994; Murie and Musterd, 1996; Van Kempen and Wissink, 2016).

In early studies on socio-spatial segregation, scholars paid attention to the spatial dimension, with particular concern about spatial concentration and the non-uniform spatial distribution of different groups of people. Scholars have pointed out that attention should also be directed toward other potentially important dimensions (White, 1983; Van Kempen, 1994; Forrest, La Grange and Yip, 2004). Research dealing with residential segregation emphasizes the role of neighborhood and residence, which is a long-standing and important theme in the literature (Morrill, 1991; Gu and Shen, 2003; Phillips, 2010; Hui et al., 2012). With the advances of technology and improved transportation services, a high mobility narrative of the current society has emerged, and the static view has drawn criticism. The argument is that social relationships are redefined with the high mobility of people today and that we should rethink the concept of neighborhood (Forrest, 2008). Our comprehensive understanding of socio-spatial segregation may be enriched further by looking into the temporal dimension and activities or social interactions outside people's home neighborhoods (Hägerstrand, 1989; Bertolini and Dijst, 2003; Sheller and Urry, 2003; Wong and Shaw, 2011; Kwan, 2013; Palmer, 2013; Wissink and van Kempen, 2016; Park and Kwan, 2018). Some concepts have been developed, such as the space–time prism (Miller, 1991), time–space trajectories (Atkinson and Flint, 2004), activity–space (Wong and Shaw, 2011), and uncertain geographic context problems (Park and Kwan, 2018) about the need to go beyond residential segregation.

This study aims to contribute to the literature and explore socio-spatial segregation by using a framework that integrates time, space, mobility, and activity components. This research measures socio-spatial segregation at the individual level and analyzes the mobility-activity patterns, highlighting differences between groups of people in terms of housing tenure status, age, and occupation. We ask the following questions: How does segregation vary in the daytime and at night? Does it change across a week and does it vary on weekdays and weekends? Are any mobility–activity patterns of people revealed by the results? How does it differ between people with different socioeconomic status, particularly in regard to variables such as housing tenure status, age, and occupation? Do any constraints exist due to home neighborhood types with regard to the exposure between people of different socioeconomic status in their non-home stations? Do we address methodological challenges in processing mobility–activity data efficiently?

To answer our research questions, a mobile phone-based activity tracking application was used to collect data. 71 participants finished data collection over a week, and their demographic data, mobility data, and activity data (with a time filter of 30 minutes) were collected simultaneously. We integrated census data and mobility-activity data within the i-STP index developed by Park and Kwan (2018). We then used decision tree algorithms, analysis of variance (ANOVA), and Tukey's honest significance difference (HSD) test to further explore the mobility–activity patterns of people. The remainder of this article is structured as follows. The next section reviews the relevant literature, followed by an introduction of the study area and the proposed methods. The measurement of segregation with i-STP index and the mobility-activity patterns identified by decision tree algorithms are then illustrated. Finally, the findings, contributions, and limitations of this research are concluded and discussed.

2. Literature review

The previous literature on socio-spatial segregation showed two main approaches, namely, place based and people based. The place-based approach focuses on home and non-home places as fixed nodes, and related discussions mainly revolve around residential segregation, socio-spatial segregation in activity places (e.g., workplace and places of leisure), the meaning of those places, and the consequent impacts (Giddens, 1984; Hägerstrand, 1985; Morrill, 1991; Reskin, 1993; Phillips, 2010). The people-based approach emphasizes people as the analyzing unit and facilitates subgroup analyses with demographic data and the comparison of the differences between people of different socioeconomic status (e.g., intergenerational, intergender, or the difference between people of different income status) (Skeggs, 1999; Farber, Páez and Morency, 2012). In the literature, place-based analysis transitions to people-based analysis, and scholars have made attempts to integrate them. Studies on socio-spatial segregation are evolving from focusing more on home neighborhoods and residential segregation to paying more attention to non-home stations, space–time paths, and mobility of people (Hägerstrand, 1985; Kwan, 2009; Wong and Shaw, 2011; Chai, 2013; Silm and Ahas, 2014a, 2014b; Li and Wang, 2016; Park and Kwan, 2018; Shoval, Schvimer and Tamir, 2018).

Recently, the analysis of socio-spatial segregation at the individual level has demonstrated promising progress, and the impacts of the practices, processes, and patterns of people's mobility are emphasized, reflecting on the new mobility narrative that the social sciences are facing today (Sheller and Urry, 2003; Farber et al., 2012; Raanan and Shoval, 2014; Silm and Ahas, 2014a; Li and Wong, 2016; Park and Kwan, 2018; Shoval, Schvimer and Tamir, 2018). The mobility of people would impact their segregation in home and non-home stations (Farber et al., 2011; Park and Kwan, 2018) and the reconstruction of social relationships outside their home neighborhoods (Jaffe, Klaufu and Colombijn, 2012; Yip, Forrest and Xian, 2016). Some scholars found that the mobility of people could alleviate residential segregation (Yip et al., 2016; Park and Kwan, 2018), and mobility appears to differ among people (Forrest, 2008; Farber et al., 2011; Jaffe et al., 2012; Yip et al., 2016). Some people are more likely than others to spend time in their home neighborhoods (Yip et al., 2016). Thus, home neighborhoods still matter, especially for people with low mobility.

As pointed out, comprehensive examinations of the full spectrum of socio-spatial segregation are still at an early stage (Hagerstrand, 1989; Kwan, 2013; Park and Kwan, 2018). The concept of multi-contextual segregation has been proposed (Park and Kwan, 2018), with an emphasis on the uneven segregation experiences of individuals within various daily life contexts. Embracing the spatial and temporal dimensions with focus on the mobility of people, Park and Kwan (2018) further developed the Individual-level Spatiotemporal Proximity Index (i-STP index), thus enabling the measurement of multi-contextual socio-spatial segregation at the individual level. By applying the index in a study in Atlanta, Georgia, they found that people experience varying levels of segregation throughout the day, with these levels generally decreasing during the daytime and increasing at night, and the differences are statistically significant. However, how does the segregation level change across a week and does it vary on weekdays and weekends? Related studies are limited (Silm and Ahas, 2014a), and further analysis is needed to provide adequate evidence.

Recently, research has paid increasing attention to information about activity types and interactions among people in analyzing socio-spatial segregation (Ellegård and Svedin, 2012; Silm and Ahas 2014a, 2014b; Yip et al., 2016). Considering the activity information in analysis is encouraged because doing so could help answer the questions "Why there?" and "With whom are those activities conducted?" The results should benefit the research theme of socio-spatial segregation that concerns the interactions and mutual activity participation of people with different socioeconomic backgrounds. However, among those studies that integrate time, space, mobility, and activity components in analysis, those measure socio-spatial segregation at the individual level are limited (e.g., Palmer, 2013; Silm and Ahas, 2014a; Yip et al., 2016; Park and Kwan, 2018). Aside from variables such as ethnic identity (Silm and Ahas, 2014a, 2014b; Raanan and Shoval, 2014), social class, or housing tenure status (Murie and Musterd, 1996; Palmer, 2013; Li and Wong, 2016), the variables of age and occupation are not adequately explored (Park and Kwan, 2018). The variable of age could relate to life cycle issues, which are important in socio-spatial segregation theories, and the variable of occupation is associated with the difference among employed people and the unemployed, such as retired elderly and housewives whose mobility may be relatively low (Rofe, 2003).

Challenges also exist in the development of methodologies for collecting and processing data efficiently. The development of GIS and GPS tracking technology has provided a promising method to simultaneously collect mobility, activity, and demographic data. With such information, richer features can be extracted compared with the previous studies (e.g., 1,006 features in this study). This situation poses a challenge to processing data efficiently in analyzing their potential mobility–activity patterns. In the previous studies, descriptive analysis, geometric analysis, and inferential statistical data analysis techniques were often employed, such as location quotient (e.g., Silm and Ahas, 2014a), multiple correspondence analysis (e.g., Yip et al., 2016), gravity-based models (e.g., Grannis, 2002), and logic regression models (Timmermans, 1996). In recent years, studies that use relatively more complicated and advanced statistical techniques have increased, such as those based on structural equation models (Ren and Kwan, 2009) or data mining models (Arentze et al., 2000). Decision tree algorithms are ideal candidates for figuring out the features that contribute the most to the separability between different groups of people to identify their potential mobility–activity patterns. As data mining approaches, decision tree algorithms have been widely used in classification applications, and they can select the optimal features that can achieve the best classification accuracy (Loh and Shih, 1997). They have been employed in studies about activity recognition (e.g., sitting, standing, walking, cycling, car-driving) (Bao and Intille, 2004), but their applications in spatiotemporal mobility-activity analysis are insufficient, with a few exceptions (Arentze et al., 2000).

3. Data and Methods

3.1 Study area

Socio-spatial segregation is a context-based question, and the importance of the local context has long been highlighted in the previous literature. Hong Kong was chosen as the study area not only because we have easy access to relevant data but also because we regard it as a typical case. The housing tenure of people is considered to be associated with their socioeconomic status (Murie and Musterd, 1996; Forrest and Lee, 2003; Forrest et al., 2004). Hong Kong has a large public housing sector, with 44.7% of its population recorded to be living in public rental housing or subsidized sale flats in 2017 (Hong Kong Statistics Department, 2018). Hence, this city is an ideal candidate to reflect on the variables of housing tenure status, namely, public rental housing, assisted housing, and private housing. Hong Kong is also an aging society and could aid the analysis of the variable of age for exploring the life cycle issues in socio-spatial segregation analysis. Moreover, according to a newly released report by UBS Prices and Earnings, Hong Kong tops global rankings in terms of weekly working hours, surpassing Mumbai, Mexico City, and New Delhi (Li, 2016). Therefore, Hong Kong is an ideal case to reflect on the variable of occupation. Although Hong Kong has a high degree of income polarization and wealth inequalities (Chiu and Liu, 2004), the residential segregation in the region is rather modest (Yip, 2012). The situation may differ if people's mobility is considered, and thus, Hong Kong serves as an ideal case for this study.

3.2 Data and data collection

To address the challenges in data collection and to improve accuracy and efficiency, this study utilized an app designed for Android mobile phones to collect the mobility, activity, and demographic data of the participants simultaneously. The app was developed specifically for this study and can be used only in Hong Kong. A later version of it was developed by the authors to also be used in Mainland China. The path data of the participants were recorded by the app every 5 minutes. If people stayed in the same location for more than 30 minutes, the participants were required to input their activity data, which involved confirming the activity locations captured automatically by GPS or typing in the true location information manually, selecting the activity type (Appendix A.1), confirming the start time and end time of the activity, and selecting with whom (self, family, friends) the activity was performed. A duration of 30 minutes was set as the filter for the participants to input their activity data during the study. This approach reduced their workload of data input and improved the percentage of participants that complete data collection. More important, this approach was selected because the authors would like to focus on those daily activities with a longer duration. Demographic data, including age, occupation, housing tenure status, and living districts, were collected through two rounds of a questionnaire survey that is built into the app. A random sampling method was used to invite participants older than 18 years to the main round of data collection. Letters were sent out by post to randomly selected addresses stratified by housing tenure status and districts in mid-2014. A total of 3,380 letters were sent out, and the participation rate was 1.33%, with 136 participants agreeing to participate and one-third of them completing the seven-day data collection. With a supplementary round of participant recruitment via snowballing, a total of 71 participants were involved in the analysis. Although the sample size was relatively small, it fit the purpose of this research, with accurate spatiotemporal data, rich activity information, and basic demographic data for the method exploration. The income data from the 2011 Hong Kong Population Census and the large street block groups (LSBG) as the neighborhood analysis unit were employed to calculate residential segregation in home neighborhoods. Each LSBG had an average population size of 4,200. These LSBGs were categorized into six types from poor to rich ones; "dumbbell" is a category characterized by a rather large proportion of poor households and rich households in terms of income (Figure 1). These home neighborhood types should help explore the constraints on the participants' exposure to people of different socioeconomic status resulting from their respective home locations.

The analysis involved 71 participants. About half of them were aged 18–30 years, and those aged 30–40 years and older than 40 years comprised a quarter of the sample. With regard to occupation, nearly one-third of the participants were students, about six in ten were employed, and the rest were housewives or retirees. For the housing tenure status, about one-third of the participants were from private housing, slightly more than one-third were from public rental housing, and the rest were from assisted housing. Although we are not aiming to represent the population with such a small sample size, caution should be observed, given that young people and those living in public housing are overrepresented; the proportion of participants living in poorer neighborhoods should likewise be considered. A variety of features were extracted from our mobility–activity data around eight aspects, namely, when, where, with whom, time length, distance, scale, ratio, and counts. Examples include the number of dinner meetings with friends, time spent on shopping, time arriving home, distance between home and workplace, ratio between time spent in shopping, and time spent in all non-routine activities. Given 16 types of activities and 3 types of accompanying persons on routine or non-routine days, a total of 1,006 features were extracted for each participant.

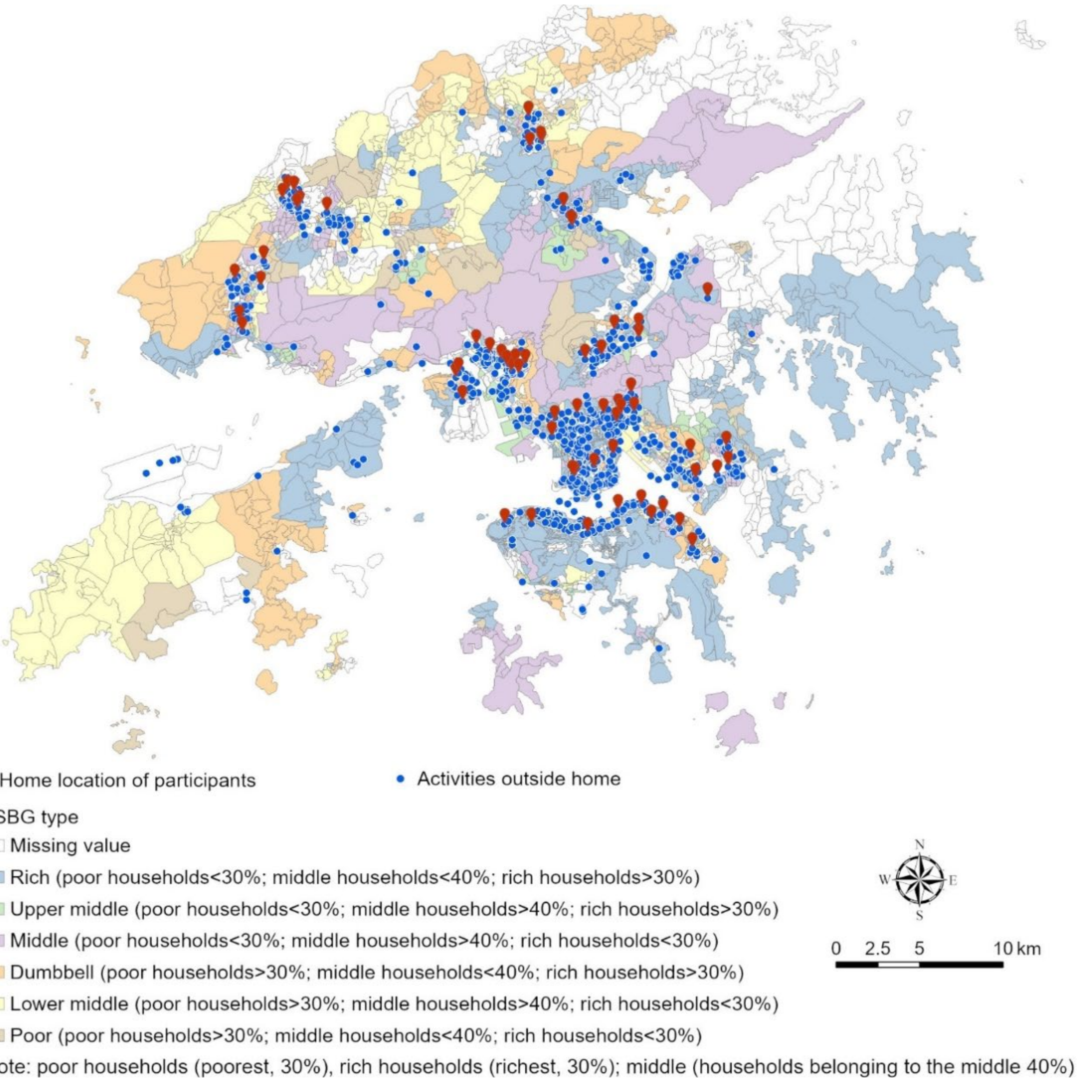


Figure 1 Study area (Hong Kong) and neighborhood types with large street block groups (LSBG)

3.3 Data processing

3.3.1 Measuring socio-spatial segregation and its temporal variations in a week with the i-STP *index*

The i-STP index was proposed by Park and Kwan (2018). It integrates the k-nearest neighbor analysis and the modified multi-group spatial proximity index of Grannis (2002). By highlighting the individual and time dimensions, the i-STP index is able to calculate the segregation index at the individual level at different times of the day. It is defined as Eq. (1).

$$i - \text{STP} = \frac{\sum_{g=1}^{n} N_g P_{gg,t}}{k * P_{kk,t}} \tag{1}$$

$$f\left(d_{P_iP_j}\right) = \exp(-2d_{p_ip_j}) \tag{2}$$

$$P_{kk,t} = \frac{1}{k^2}\sum_i \sum_j f(d_{P_iP_j}) \tag{3}$$

$$P_{gg,t} = \frac{1}{{N_g}^2}\sum_i \sum_j f(d_{P_iP_j}) \tag{4}$$

where *i*-STP index refers to the individual i's spatial proximity index value during a specific time period t (e.g., t = 1 represents the time period 12 am–3 am). k is a predefined number of the nearest neighbors in the given time period and could be adjusted. Defined as Eqs. (2) and (3), respectively, $P_{kk,t}$ is the average proximity between all k neighbors during t, while $P_{gg,t}$ is the average proximity between individuals of g group among k neighbors during t. N_g is the mean number of individuals of g group among k neighbors (the region-wide proportion of g group multiplied by k).The proximity between two individuals p_i and p_j is defined as Eq. (4), where $d_{p_ip_j}$ is the distance between the two individuals. Mainly considering that the locations a person passes by when moving may be too fleeting to trigger segregation experience, the *i*-STP index focuses on activity locations only. For an individual who conducts more than one activity during the given time period, the segregation levels are measured separately for each activity first and then average weighted by its time duration to generate a single segregation index value.

By applying the *i*-STP index developed by Park and Kwan (2018) and on the basis of the data collected in Hong Kong, this study measured the socio-spatial segregation index at the individual level and examined its dynamic changes across a week. Park and Kwan (2018) found that the change in k and that in the distance threshold of k-nearest neighbors do not significantly affect the segregation analysis results. In the present study, different values of k were employed to explore the sensitivity and credibility issue. We adopted $k = 6$, $k = 10$, and $k = 15$ to test the sensitivity and credibility of the results, and no distance threshold is employed mainly in consideration of the small sample size of 71 participants. The analysis found that the change in k value does not significantly affect the general trends (refer to the sensitivity analysis in Appendix A.3), and then $k = 6$ is selected for the calculation and presentation of results. The temporal unit was set to 3 hours, which is the same as that in Park and Kwan's research (2018). This value could also be adjusted to a small analysis unit. To explore the potential differences between the socio-spatial segregation level during the day and at night, we calculated the mean values of the *i*-STP index of an individual from 6 am to 6 pm and that from 6 pm to 6 am by approximating the socio-spatial segregation level during the day and at night, respectively. T-test was then performed to check whether the difference in the *i*-STP index between the daytime and at night and that between weekdays and weekends is significant. Different from Park and Kwan's calculation (2018) based on weekday data, the current research further extended the measurement period to a week, such that weekdays and weekends were included, to provide further insights into the consistency of potential patterns of socio-spatial segregation levels.

With the method proposed by Park and Kwan (2018) and the value of the *i*-STP index obtained in the last step, the residential segregation of an individual was then calculated by using the mean of the *i*-STP value of the same person when he or she stayed in their home neighborhood. The multi-contextual segregation level was calculated by the average *i*-STP value of each individual inside and outside the home neighborhood. Then, T-test was performed to explore potential differences between multi-contextual segregation and residential segregation at the individual level. Further subgroup analysis in terms of housing tenure status, home neighborhood types, age, and occupation was conducted.

3.3.2 Exploring mobility–activity patterns using decision tree algorithms

Using the mobility–activity data collected in Hong Kong, this study proposed to explore the mobility–activity patterns of the participants on the basis of the classification rules provided by the decision trees and the results of supplemented ANOVA and Tukey's HSD test. As discussed, although the sample size is small, rich features of each participant—1,006 features in this study—were extracted from the accurate spatiotemporal data, rich activity information, and basic demographic data. This situation poses a challenge to processing data efficiently in analyzing the potential mobility–activity patterns of people. Decision tree algorithms were employed because they can select the optimal features, in which different groups of people had the highest separability, from the numerous features extracted from the mobility–activity data. Moreover, the classification rules obtained by decision tree algorithms are easy to understand and interpret, helping figure out the mobility-activity patterns of people. Moreover, decision trees require little data preparation and can handle various types of features (e.g., categorical and numerical features), and they are robust and efficient when dealing with large datasets. The decision tree algorithms split the samples into two parts in each node, during which all the features are evaluated one by one and the optimal one that results in the most homogeneous subnodes is selected. Such binary splitting is implemented on the resultant subnodes recursively, and then the features in which different groups exhibit the biggest differences are identified. Decision tree algorithms do not include validation in the algorithms themselves; thus, ANOVA and Tukey's HSD test were performed to further examine whether the separability in the selected features was statistically significant. Therefore, decision tree algorithms supplemented by ANOVA and Tukey's HSD test were proposed by this study as ideal candidates for the exploration of mobility-activity patterns. In this research, the decision tree algorithm embedded in the R package was employed.

4. Results

4.1 Beyond residential segregation: decreased socio-spatial segregation levels during the daytime and a consistent trend in a week

As shown in the results of the *i*-STP index, the socio-spatial segregation levels vary across a day and across a week (Figure 2). In general, the mean values of socio-spatial segregation levels decrease in daytime and increase at night. This trend is reflected both in the results of *i*-STP calculated by housing tenure types and that calculated by home neighborhood types (Appendix A.2). For the results based on housing tenure status, T-test is further conducted, and significant differences are found between the segregation level in the daytime from 6 am to 6 pm and that at night from 6 pm to 6 am (Table 1 and Figure 3). This result verifies the finding by Park and Kwan (2018) based on weekday data on Atlanta, Georgia, USA, as well as the outcomes of previous studies (e.g., Silm and Ahas, 2014a; Le Roux, Vallée and Commenges, 2017). Unlike the previous studies, the present research extended the study period to a week, and a consistent trend is observed in the results (Figure 2). However, no significant difference is found in the further subgroup analysis in terms of housing tenure status, home neighborhood types, age, and occupation. Moreover, socio-spatial segregation levels are generally higher on the weekends than on the weekdays both in the daytime and at night, but the difference is not statistically significant (Table 1).

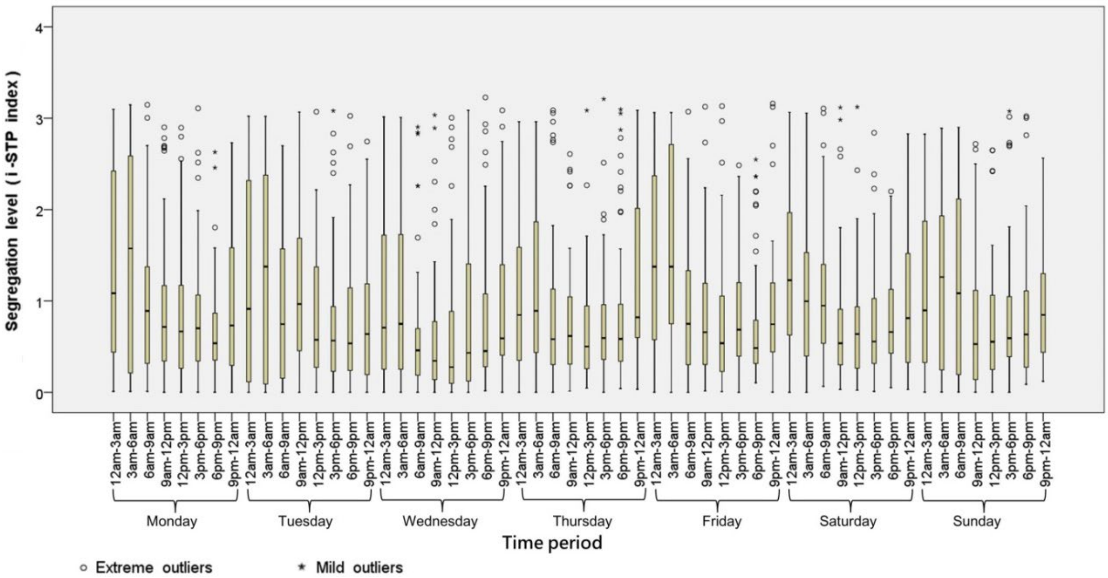


Figure 2 Temporal variations in segregation levels by housing tenure status in a week

Table 1 Paired sample T-test (Tukey's test)

Difference of levels	Difference of means	*P*-value	Significant
Day–night	−.2004306	**.000**	**Yes**
Multi-contextual segregation–residential segregation	−.2026760	**.002**	**Yes**
Daytime weekday–daytime weekend	−.0316033	.672	No
Night weekday–night weekend	−.0564347	.486	No

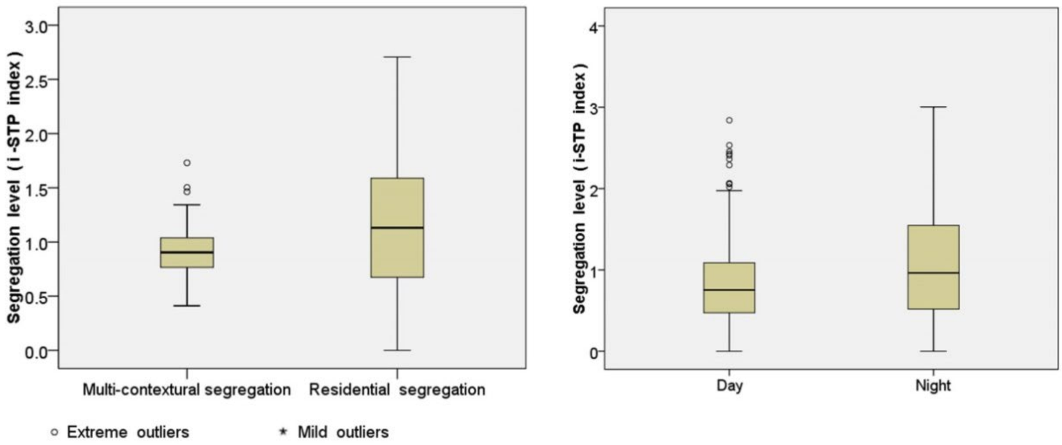


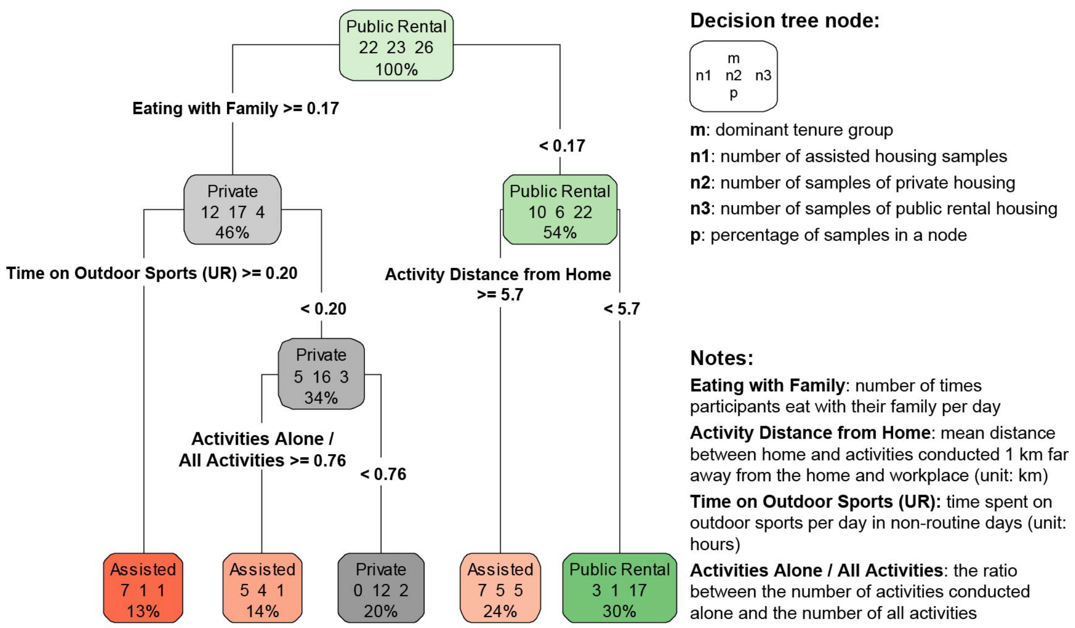
Figure 3 Differences in multi-contextual segregation and residential segregation during the day and at night

The mean value of multi-contextual segregation and that of residential segregation level at the individual level are then calculated and compared. The mean value of multi-contextual segregation is significantly lower than that of residential segregation (Table 1). The results give empirical support to the proposition that the mobility of people alleviates their residential segregation. However, no significant difference is found in further subgroup analysis between people of different housing types, home neighborhood types, age, and occupation. This finding means that the same index variability trend applies to all the identified groups and that no significant subgroup differences exist.

4.2 Mobility–activity patterns and constraints due to home location

The proposed method using decision tree algorithms supplemented by ANOVA and Tukey's HSD test was effective and efficient in identifying the mobility-activity patterns of participants, and some significant results are found. For the classification of housing tenure status, the finding that people from private housing dine with their families more often than those from public rental housing do is statistically significant. People from assisted housing tend to travel longer distances from home to conduct activities than those from public rental housing do. People from public rental housing tend to keep to themselves more than those from private housing do (Figure 4).

(a)



(b)

	Eating with Family		Activity Distance from Home		Acting Alone / All Activities	
	P-value	Sig.	P-value	Sig.	P-value	Sig.
Private and Assisted	0.83	No	0.42	No	0.38	No
Public rental and Assisted	0.10	No	**0.03**	**Yes**	0.27	No
Public rental and Private	**0.02**	**Yes**	0.39	No	**0.01**	**Yes**

(c)

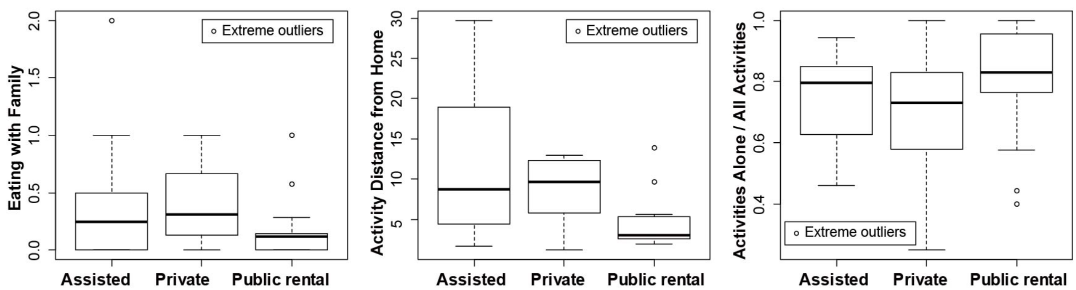
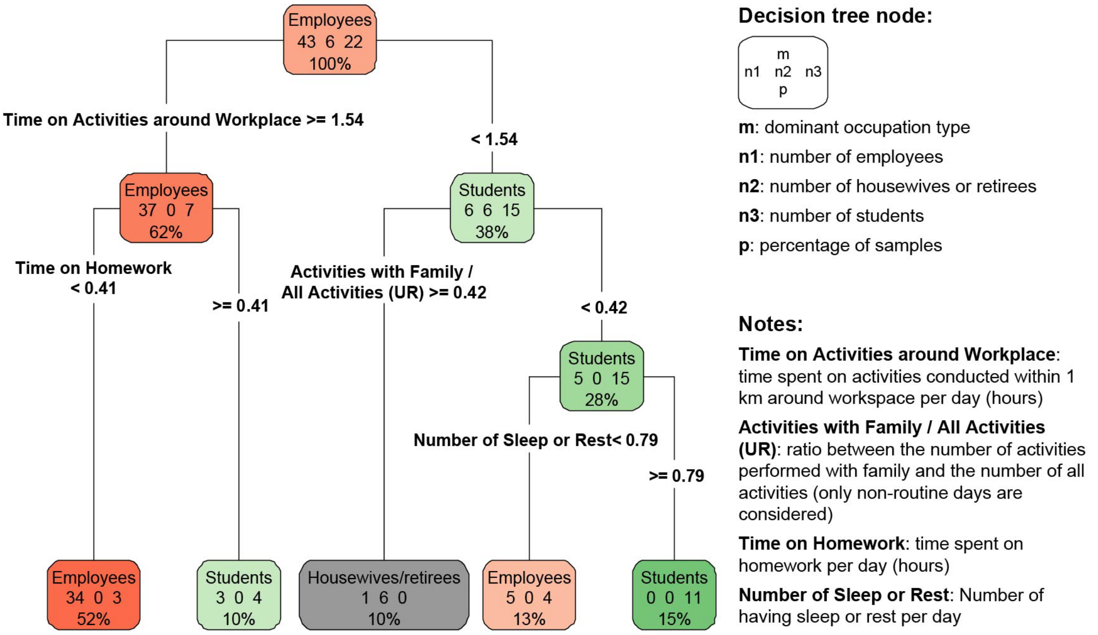


Figure 4 Classification results of people of different housing tenure types (a) Decision tree for classifying different housing tenures, (b) Significance tests of features selected by the decision tree, (c) Variation in the features between different housing tenures

In terms of occupation, students show a lower ratio between the number of activities performed with family and the total number of activities in non-routine days in comparison with housewives, retirees, and employed people (Figure 5).

(a)



(b)

	Time on Activities around Workplace		Activities with Family / All Activities (UR)		Number of Sleep or Rest	
	P-value	Sig.	P-value	Sig.	P-value	Sig.
Students-Housewives/retirees	0.34	No	**0.00**	**Yes**	0.12	No
Employees-Housewives/retirees	**0.00**	**Yes**	**0.01**	**Yes**	0.72	No
Employees-Students	**0.00**	**Yes**	**0.00**	**Yes**	0.07	No

(c)

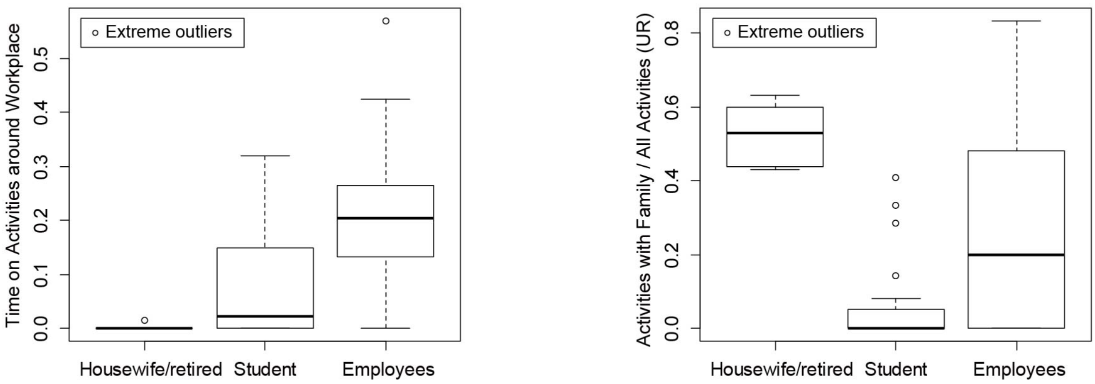
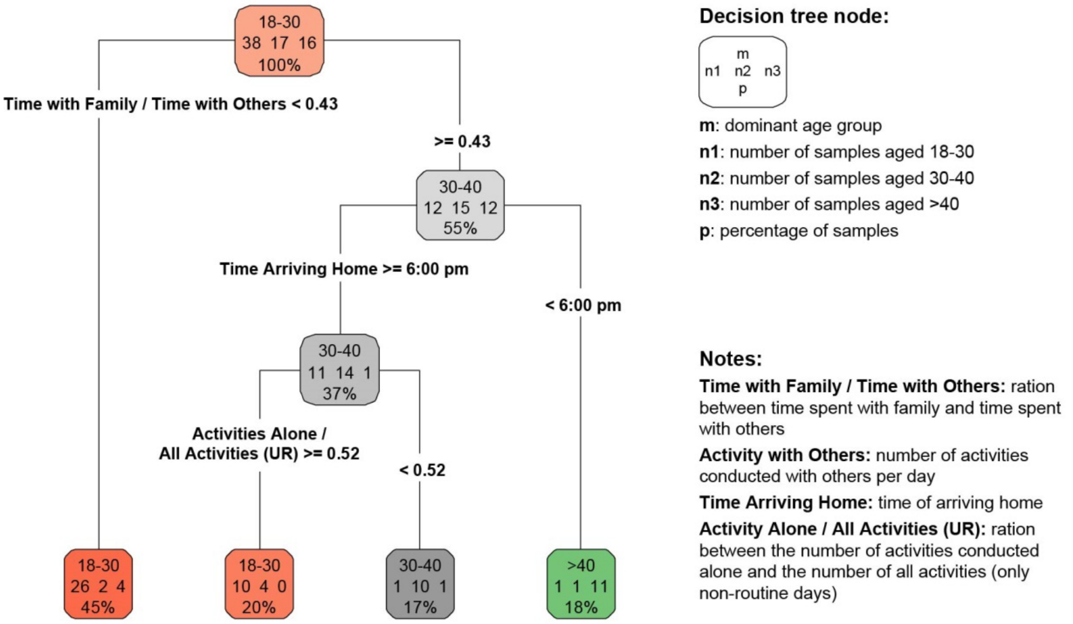


Figure 5 Classification results of different occupation groups (a) Decision tree for classifying different occupation groups, (b) Significance tests of features selected by the decision tree, (c) Variation in the features between different occupation groups

The same finding is also supported by the classification results of different age groups (Figure 6), that is, compared with older participants, young people aged 18 to 30 years feature a smaller ratio between activity time with family and that with others. This result may indicate that unlike young people, older people place great value on activities with the family. Significant differences are also found between people aged more than 40 years and those aged 30–40 years in terms of the time of arriving home. People older than 40 years return home early, whereas those between 30 and

40 years tend to return home later than 6 pm.

(a)



(b)

	Time with Family / Time with Others		Time Arriving Home	
	P-value	Sig.	P-value	Sig.
Age 18-30 vs. Age 30-40	**0.00**	**Yes**	0.50	No
Age 18-30 vs. Age >40	**0.00**	**Yes**	0.17	No
Age 30-40 vs. Age >40	0.54	No	**0.04**	**Yes**

(c)

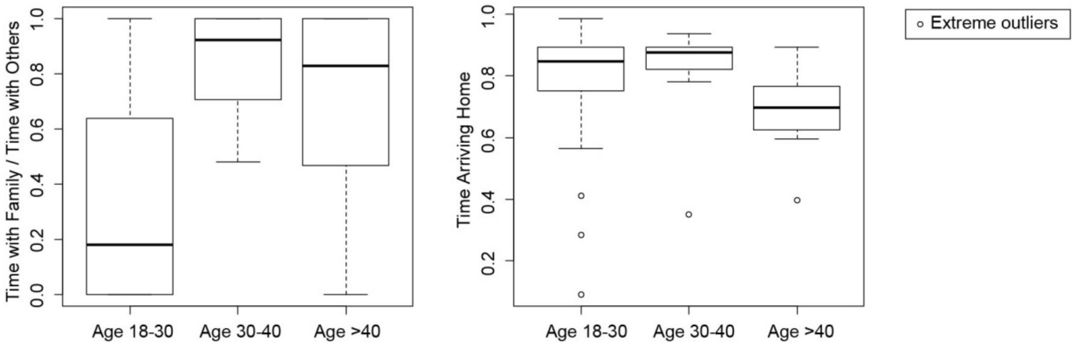
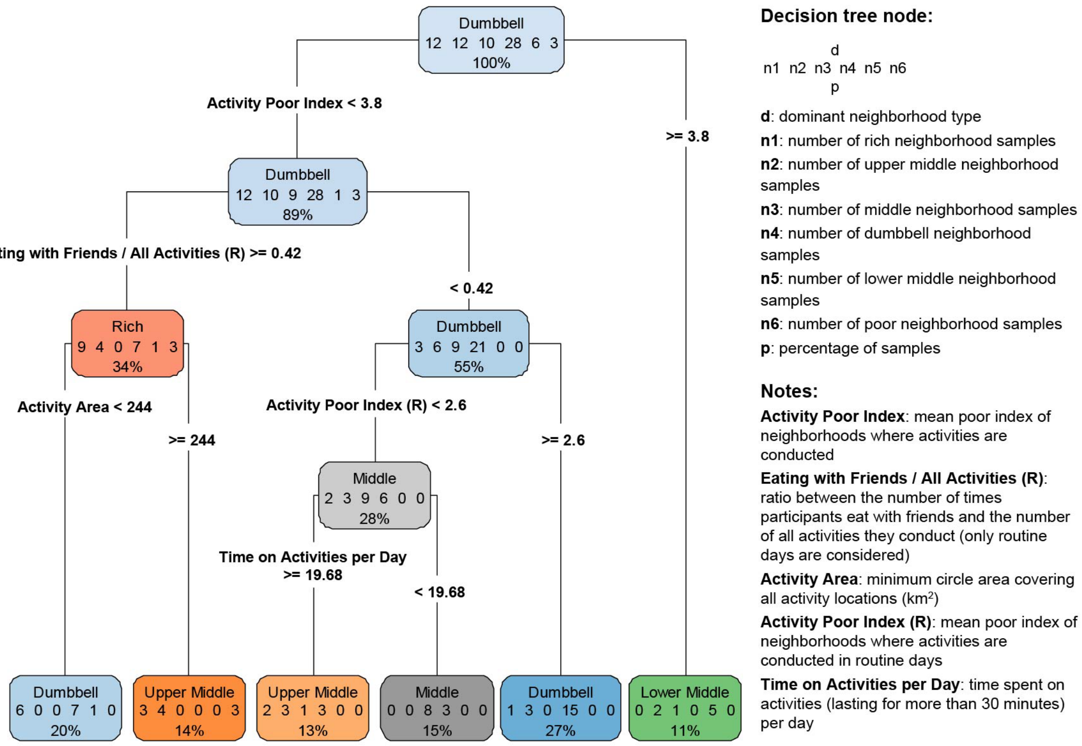


Figure 6 Classification results of different age groups (a) Decision tree for classifying different age groups, (b) Significance tests of features selected by the decision tree, (c) Variation in the features between different age groups

We are also interested in the constraints brought by our participants' home neighborhoods to their exposure in non-home stations to people of different social status. Thus, aside from housing tenure status, occupation, and age, the classification of home neighborhood types varying from poor to rich is analyzed further. As shown in Figure 1, the home neighborhoods were categorized into six types, namely, poor, lower middle, dumbbell, middle, upper middle, and rich. "Dumbbell" means the neighborhood has a rather large proportion of poor households and rich households in terms of income. Compared with people from middle and dumbbell neighborhoods, those from rich neighborhoods dine with their friends more often on their routine days. More importantly, participants from lower middle neighborhoods tend to conduct more activities in poorer neighborhoods than those from dumbbell, middle, upper middle, and rich neighborhoods both on routine days and non-routine days. A similar difference is found between those from rich and dumbbell neighborhoods. For activities on routine days, significant differences are observed between our participants from lower middle neighborhoods and those from middle and rich neighborhoods. This finding means that people from poorer neighborhoods tend to conduct activities in poorer neighborhoods (Figure 7).

(a)



(b)

	Activity Poor Index		Activity Poor Index (R)		Eating with Friends / All Activities (R)	
	P-value	Sig.	P-value	Sig.	P-value	Sig.
Lower Middle-Dumbbell	**0.01**	**Yes**	0.21	No	0.98	No
Middle-Dumbbell	0.90	No	0.39	No	0.79	No
Poor-Dumbbell	1.00	No	0.95	No	0.08	No
Rich-Dumbbell	**0.04**	**Yes**	0.30	No	**0.02**	**Yes**
Upper Middle-Dumbbell	0.99	No	1.00	No	0.50	No
Middle-Lower Middle	**0.00**	**Yes**	**0.01**	**Yes**	1.00	No
Poor-Lower Middle	0.42	No	0.98	No	0.08	No
Rich-Lower Middle	**0.00**	**Yes**	**0.01**	**Yes**	0.06	No
Upper Middle-Lower Middle	**0.01**	**Yes**	0.44	No	0.49	No
Poor-Middle	0.90	No	0.41	No	**0.02**	**Yes**
Rich-Middle	0.62	No	1.00	No	**0.01**	**Yes**
Upper Middle-Middle	1.00	No	0.40	No	0.16	No
Rich-Poor	0.30	No	0.38	No	1.00	No
Upper Middle-Poor	0.96	No	0.99	No	0.60	No
Upper Middle-Rich	0.33	No	0.34	No	0.66	No

(c)

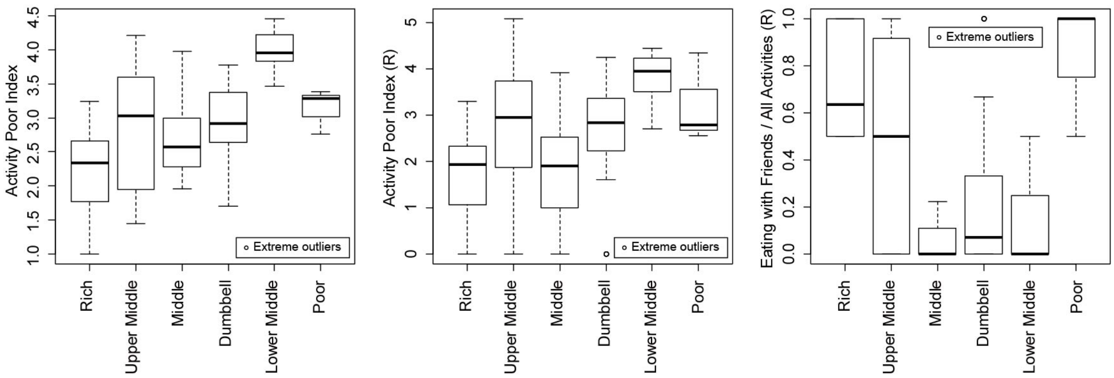


Figure 7 Classification results of people of different home neighborhood types (a) Decision tree for classifying different home neighborhood types, (b) Significance tests of features selected by the decision tree, (c) Variation in the features between different home neighborhood types

5. Discussion and conclusions

On the basis of a case study of Hong Kong and a mobility–activity dataset collected by a mobile phone app, this study explores socio-spatial segregation under a framework that embraces spatial and temporal dimensions, with emphasis on the mobility and activity of people. Attempts have been made to go beyond residential segregation, and special attention has been paid to the interactions among people of different socioeconomic status and the variables of housing tenure status, age, and occupation. With the use of the *i*-STP index developed by Park and Kwan (2018), the segregation index at the individual level is calculated and the measurement period is extended to a week compared with previous studies. The results show that the socio-spatial segregation level generally decreases in the daytime and increases at night, with the difference being statistically significant. This outcome verifies the findings by Park and Kwan (2018) using weekday data on Atlanta, Georgia, USA. A consistent trend across the week, along with a few variations on weekends, is further identified by extending the study period to one week. Furthermore, the mean value of multi-contextual segregation is significantly lower than that of residential segregation. The results of the present study support the argument that social relationships are redefined with the high mobility of people today (Forrest, 2008). Similar to the findings of Silm and Ahas (2014a), Li and Wong (2016), and Park and Kwan (2018), empirical evidence that mobility may alleviate residential segregation was provided by the present study. However, no significant differences are found between subgroups in terms of age, occupation, housing types and home neighborhood types.

In the concept of socio-spatial segregation, interactions and mutual activity participation of people with different socioeconomic background are emphasized. To better understand people's interaction and segregation with special attention to their different socioeconomic status, decision tree analysis focusing on the mobility-activity patterns of different people was conducted, and some significant subgroup differences were found. The results indicate some social norms, including dining with family or friends, eating at home, time of arriving home, young people or students being together with family, and some people hanging out alone. These findings help answer the questions "Why there?" and "With whom are certain activities conducted?" Moreover, according to the results of the decision tree algorithms and the subsequent difference tests for the variable of home neighborhood types, the home locations of our participants result in constraints on their exposure in non-home stations to people of different socioeconomic status. That is, people from poorer neighborhoods tend to conduct activities in poorer neighborhoods. Although a significant difference is found with the i-STP index between multi-contextual segregation and residential segregation, empirical support is provided in the decision tree results for the possible connections between people's residential segregation and their segregation that goes beyond home neighborhoods. More important, such connections vary among different groups of people. Scholars (Giddens, 1984) argue that constraints could also be viewed as possibilities, and we see these complications as opportunities to obtain a better understanding of the mechanism.

Attempts have been made by this study to improve the robustness of the analysis. A random sampling method was used to invite participants to the main round of data collection, and ANOVA and Tukey's HSD test with significance test were used to validate the results by decision tree algorithms. However, as discussed, we do not aim to represent the population with a small sample size of 71 participants, and precautions may need to be taken against possible sample bias. Although the app-based data collection method does bring advantages, including simultaneous and accurate data collection, it could also contribute to sample bias if people who do not use smartphones or Android mobile phones are excluded. A duration of 30 minutes were set as the filter for the participants to input their activity data during the study. Admittedly, activities with a short duration might have been omitted, and this omission could also contribute to sample bias.

Despite those limitations, this study identified several topics of concern. A primary concern is the theoretical approach. An integral approach that embraces spatial, temporal, mobility, and activity dimensions is valuable for a richer understanding of socio-spatial segregation that goes beyond home neighborhoods. Paying special attention to the microscale, various time resolutions, and interactions among people of different socioeconomic status also helps. Another concern is related to the proposed method. In exploring the temporal variation of socio-spatial segregation beyond home neighborhoods, the use of a mobile phone app in data collection and the employment of the i-STP index and decision tree algorithms supplemented by ANOVA and Tukey’s HSD test in data processing are effective and efficient. This approach allows segregation to be measured at the individual level; extends the measurement period to test the consistency of potential patterns; highlights concerns about mobility and activity, particularly the interactions of people with different socioeconomic backgrounds; and addresses challenges to processing data efficiently. In this regard, both the method and findings of this study may provide implications to the literature on socio-spatial segregation and for policymakers working toward a more equitable city.

References

Arentze, T. A., Hofman, F., Van Mourik, H., Timmermans, H. J., & Wets, G. (2000). Using Decision Tree Induction Systems for Modeling Space‐Time Behavior. *Geographical analysis*, *32*(4), 330-350.

Atkinson, R., & Flint, J. (2004). Fortress UK? Gated communities, the spatial revolt of the elites and time–space trajectories of segregation. *Housing studies*, *19*(6), 875-892.

Bao, L., & Intille, S. S. (2004, April). Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing* (pp. 1-17). Springer, Berlin, Heidelberg.

Bertolini, L., & Dijst, M. (2003). Mobility environments and network cities. *Journal of Urban Design, 8*, 27-43.

Chai, Y. (2013). Space–time behavior research in China: Recent development and future prospect: Space–time integration in geography and GIScience. *Annals of the Association of American Geographers*, *103*(5), 1093-1099.

Chiu, S. W., & Lui, T. L. (2004). Testing the global city-social polarisation thesis: Hong Kong since the 1990s. *Urban Studies, 41*, 1863-1888.

Ellegård, K., & Svedin, U. (2012). Torsten Hägerstrand's time-geography as the cradle of the activity approach in transport geography. *Journal of Transport Geography*, *23*, 17-25.

Farber, S., Páez, A., & Morency, C. (2012). Activity spaces and the measurement of clustering and exposure: a case study of linguistic groups in Montreal. *Environment and Planning-Part A, 44*, 315-332.

Forrest, R. (2008). Who cares about neighbourhoods? *International Social Science Journal, 59*, 129-141.

Forrest, R., La Grange, A., & Yip, N. M. (2004). Hong Kong as a global city? Social distance and spatial differentiation. *Urban Studies*, 41(1), 207-227.

Forrest, R., & Lee, J. (Eds.). (2003). *Housing and social change: East-West perspectives*. Routledge.

Giddens, A. (1984). The constitution of society: Outline of the theory of structure. Berkeley.

Grannis, R. (2002). Discussion: segregation indices and their functional inputs. *Sociological methodology*, *32*(1), 69-84.

Gu, C., & Shen, J. (2003). Transformation of urban socio-spatial structure in socialist market economies: the case of Beijing. *Habitat International*, *27*(1), 107-122.

Hägerstrand, T. (1985). Time-geography: focus on the corporeality of man, society, and environment. *The science and praxis of complexity*, 193-216.

Hägerstrand, T. (1989, December). Reflections on "what about people in regional science?". In *Papers of the regional science association* (Vol. 66, No. 1, pp. 1-6). Springer-Verlag.

Hong Kong Statistics Department (2018), *Housing in Figures 2018*, https://www.thb.gov.hk/eng/psp/publications/housing/HIF2018.pdf, accessed on 25 Oct 2018.

Hui, E. C., Li, S. M., Wong, F. K., Yi, Z., & Yu, K. H. (2012). Ethnicity, cultural disparity and residential mobility: Empirical analysis of Hong Kong. *Habitat International*, *36*(1), 1-10.

Jaffe, R., Klaufus, C., & Colombijn, F. (2012). Mobilities and mobilizations of the urban poor. *International Journal of Urban and Regional Research, 36*, 643-654.

Kwan, M. P. (2009). From place-based to people-based exposure measures. *Social Science &*

Medicine, 69, 1311-1313.

Kwan, M. P. (2013). Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility: Space–time integration in geography and GIScience. *Annals of the Association of American Geographers*, *103*(5), 1078-1086.

Le Roux, G., Vallée, J., & Commenges, H. (2017). Social segregation around the clock in the Paris region (France). *Journal of Transport Geography*, *59*, 134-145.

Li, F., & Wang, D. (2016). Measuring urban segregation based on individuals' daily activity patterns: A multidimensional approach. *Environment and Planning A: Economy and Space*, *49*(2), 467-486.

Li, S. (2016). Hong Kong has longest working week of 71 cities. *China Daily Asia*. https://www.chinadailyasia.com/hknews/2016-05/25/content_15439024.html, accessed on 22 Jan 2022

Lieberson, S., & Carter, D. K. (1982). A model for inferring the voluntary and involuntary causes of residential segregation. Demography, 19(4), 511-526.

Loh, W. Y., & Shih, Y. S. (1997). Split selection methods for classification trees. *Statistica sinica*, 815-840.

Miller, H. J. (1991). Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information System*, *5*(3), 287-301.

Morrill, R. L. (1991). On the measure of geographic segregation. *Geography research forum*, 11, 25-36.

Murie, A., & Musterd, S. (1996). Social segregation, housing tenure and social change in Dutch cities in the late 1980s. *Urban studies*, *33*(3), 495-516.

Palmer, J. R. (2013). *Activity-space segregation: Understanding social divisions in space and time*, Doctoral dissertation, Princeton University.

Park, R. E., Burgess, E. W., & McKenzie, R. D. (1925). *The city*. University of Chicago Press.

Park, Y. M., Kwan, M. P (2018). Beyond residential segregation: A spatiotemporal approach to examining multi-contextual segregation. *Computers, Environment and Urban Systems*, 71, 98-108.

Phillips, D. (2010). Minority Ethnic Segregation, Integration and Citizenship: A European Perspective. *Journal of Ethnic and Migration Studies, 36*, 209-225.

Raanan, M. G., & Shoval, N. (2014). Mental maps compared to actual spatial behavior using GPS data: A new method for investigating segregation in cities. *Cities*, *36*, 28-40.

Ren, F., & Kwan, M. P. (2009). The impact of the Internet on human activity–travel patterns: analysis of gender differences using multi-group structural equation models. *Journal of Transport Geography*, *17*(6), 440-450.

Reskin, B. (1993). Sex segregation in the workplace. *Annual review of sociology*, *19*(1), 241-270.

Rofe, M. W. (2003). 'I want to be global': Theorising the gentrifying class as an emergent élite global community. *Urban Studies, 40*, 2511-2526.

Sheller, M., & Urry, J. (2003). Mobile transformations of public and private life. *Theory, Culture & Society, 20*, 107-125.

Shoval, N., Schvimer, Y., & Tamir, M. (2018). Tracking technologies and urban analysis: Adding the emotional dimension. *Cities*, *72*, 34-42.

Silm, S. , & Ahas, R. (2014a). The temporal variation of ethnic segregation in a city: evidence from a mobile phone use dataset. Social Science Research, 47, 30-43.

Silm, S., & Ahas, R. (2014b). Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Annals of the Association of American Geographers*, *104*(3), 542-559.

Skeggs, B. (1999). Matter out of place: Visibility and sexualities in leisure spaces. *Leisure studies*, *18*(3), 213-232.

Timmermans, H. J. (1996). A stated choice model of sequential mode and destination choice behaviour for shopping trips. *Environment and Planning A*, *28*(1), 173-184.

Van Kempen, E. T. (1994). The dual city and the poor: social polarisation, social segregation and life chances. *Urban studies*, *31*(7), 995-1015.

White, M. J. (1983). The measurement of spatial segregation. *American journal of sociology*, 88(5), 1008-1018.

Wissink, B., Schwanen, T., & van Kempen, R. (2016). Beyond residential segregation: Introduction. *Cities*, 59, 126-130. doi.org/10.1016/j.cities.2016.08.010.

Wong, D. W., & Shaw, S. L. (2011). Measuring segregation: An activity space approach. *Journal of geographical systems*, *13*(2), 127-145.

Yip, N. M. (2012). Residential Segregation in an Unequal City: Why are there Urban Ghettos in Hong Kong? In T. Maloutas & K. Fujita (Eds.), *Residential Segregation in Comparative Perspective*. London: Ashgate.

Yip, N. M., Forrest, R., & Xian, S. (2016). Exploring segregation and mobilities: Application of an activity tracking app on mobile phone. *Cities*, *59*, 156-163.

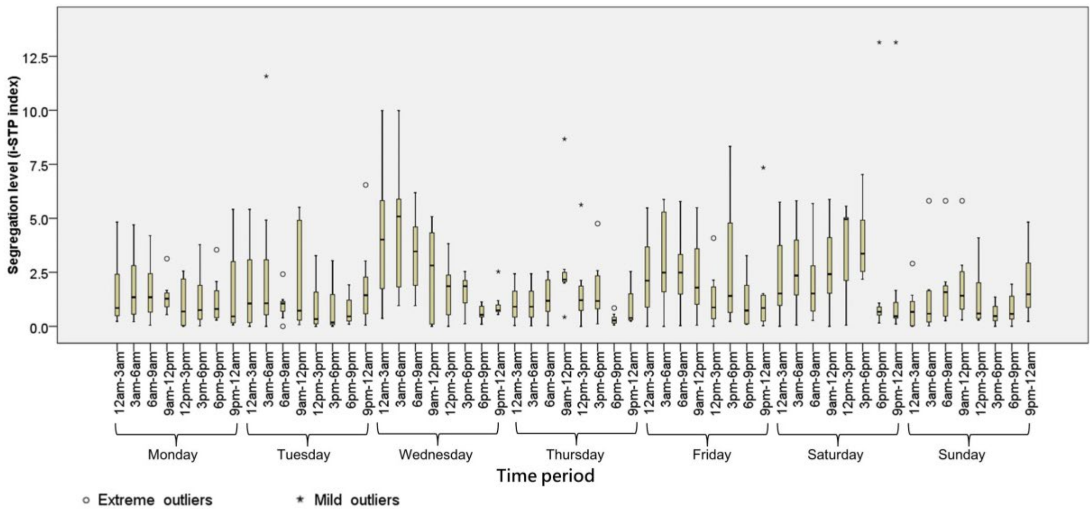
Yuan, D. (1963). Voluntary Segregation: A Study of New Chinatown. *Phylon (1960-), 24*(3), 255-265. doi:10.2307/273399.

Appendices

Appendix A.1 Activity types

Activity Types	
1.Rest and sleep	9.Waiting for people
2.Study (class/fieldtrip/lab)	10.Visits (friends/relatives) *
3.Have meals (lunch/dinner/tea time) *	11.Church
4.Go to work/Part time job	12.Housework*
5.Shopping*	13.Doctor*
6.Indoor exercise	14.Volunteer*
7.Entertainment (e.g. movie/show) *	15.Transportation
8.Outdoor activities (e.g. sports/hiking/picnic) *	16. Others
	** Indicates the requirement of the input of accompanying persons*

Appendix A.2 Temporal variations in segregation levels by home neighborhood types in a week



Appendix A.3 Temporal variations in segregation levels by housing tenure status in a week and sensitivity analysis with different k-scales including $k = 6$, $k = 10$, and $k = 15$

Difference of levels	k=6	k=10	k=15
	Significant	**Significant**	**Significant**
Day – Night	**Yes**	**Yes**	**Yes**
Multi-contextual segregation –Residential segregation	**Yes**	**Yes**	**Yes**
Daytime weekday— Daytime weekend	No	No	No
Night weekday – Night weekend	No	No	No