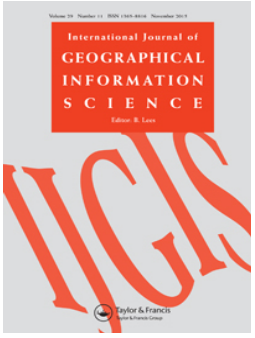

Taylor & Francis
Taylor & Francis Group

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: http://www.tandfonline.com/loi/tgis20

Everyday space–time geographies: using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn

R. Ahas, A. Aasa, Y. Yuan, M. Raubal, Z. Smoreda, Y. Liu, C. Ziemlicki, M. Tiru & M. Zook

To cite this article: R. Ahas, A. Aasa, Y. Yuan, M. Raubal, Z. Smoreda, Y. Liu, C. Ziemlicki, M. Tiru & M. Zook (2015) Everyday space–time geographies: using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn, International Journal of Geographical Information Science, 29:11, 2017-2039, DOI: 10.1080/13658816.2015.1063151

To link to this article: http://dx.doi.org/10.1080/13658816.2015.1063151

Published online: 24 Jul 2015.

International Journal of Geographical Information Science, 2015
Vol. 29, No. 11, 2017–2039, http://dx.doi.org/10.1080/13658816.2015.1063151

Everyday space–time geographies: using mobile phone-based sensor data to monitor urban activity in Harbin, Paris, and Tallinn

R. Ahas[a,b]*, A. Aasa[a], Y. Yuan[c], M. Raubal[d], Z. Smoreda[e], Y. Liu[f], C. Ziemlicki[c], M. Tiru[a,g] and M. Zook[h]

[a]*Department of Geography, University of Tartu, Tartu, Estonia;* [b]*Department of Geography, Ghent University, Ghent, Belgium;* [c]*Department of Geography, Texas State University - San Marcos, San Marcos, USA;* [d]*Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland;* [e]*Sociology and Economics of Networks and Services department/Orange Labs, Paris, France;* [f]*Institute of Remote Sensing and Geographical Information Systems, Peking University, Beijing, China;* [g]*Positium LBS, Tartu, Estonia;* [h]*Department of Geography, University of Kentucky, Lexington, USA*

(*Received 19 November 2014; accepted 2 June 2015*)

This paper proposes a methodology for using mobile telephone-based sensor data for detecting spatial and temporal differences in everyday activities in cities. Mobile telephone-based sensor data has great applicability in developing urban monitoring tools and smart city solutions. The paper outlines methods for delineating indicator points of temporal events referenced as 'midnight', 'morning start', 'midday', and 'duration of day', which represent the mobile telephone usage of residents (what we call social time) rather than solar or standard time. Density maps by time quartiles were also utilized to test the versatility of this methodology and to analyze the spatial differences in cities. The methodology was tested with data from cities of Harbin (China), Paris (France), and Tallinn (Estonia). Results show that the developed methods have potential for measuring the distribution of temporal activities in cities and monitoring urban changes with georeferenced mobile phone data.

Keywords: time use; smart city; social time; urban; mobile positioning; geography; spatial mobility

1. Introduction

Cities have always been dynamic locales and current policies associated with smart city technologies have introduced demand for monitoring tools that can quickly detect changes and identify urban rhythms (Griffiths *et al*. 2010, Batty *et al*. 2012). For example, such information is useful for better timing transportation services and ultimately developing intelligent transportation systems; for managing the operating hours for public and private services; and for developing dynamic taxation systems based on temporally variable demand, such as congestion pricing for roads or parking. There are a growing number of smart city solutions that use sensor data from various social systems (Hancke *et al*. 2013), which contrast with more traditional, static data – population registers, land-use data, questionnaires – hitherto used in urban studies. Recent approaches to city governance based on smart city solutions have sought more dynamic indicators and data sources to monitor daily life and short-term processes (Lee and Lee 2014). Such datasets are

*Corresponding author. Email: rein.ahas@ut.ee

derived from transportation and energy networks, pedestrian sensors, communication networks, and many other sources.

Mobile phone-based data are particularly promising sources for monitoring dynamic processes in cities as penetration rates for phones are high, and by design they are nearly always at hand. Several use cases have demonstrated the usefulness of using mobile data for detecting or measuring changes in urban space, particularly for monitoring purposes that rely on fast data collection (Ahas *et al*. 2007, Calabrese *et al*. 2011, Csáji *et al*. 2013, Silm and Ahas 2014, Yuan and Raubal 2014). The most common use for mobile data is mapping *de facto* distributions of population groups, spatial mobility, and social networks (Lambiotte *et al*. 2008, Sobolevsky *et al*. 2013), as well as monitoring temporal distribution of activities.

The objective of this article is to develop a methodology that uses call detail record (CDR) data of mobile network operators (MNOs) to detect spatial and temporal differences in everyday activities, or what we refer to as *social time* to distinguish it from standard or solar time. Our research questions are:

(1) How can we define and calculate meaningful temporal indicators within social time using CDR data?
(2) How can we measure and visualize the spatial differences of these indicators of social time?
(3) How do these social time indicators vary between Chinese, Estonian, and French cities?

Pursuant to the concept of the studies of cyclic processes, we define four social time indicators based on diurnal call activity curves and analyze the spatial differences of these indicators on the basis of four temporal quartiles to better understand the temporal dimension of urban life. Note that the defined indicators arise from various descriptive statistics and critical points of call activity curves (e.g., global/local maximum, derivatives, and quantiles). Because the shape of the curves may vary for different cities/countries, here, we do not attempt to cover all potential indicators, instead, we aim to demonstrate the effectiveness of employing the 'social time' concept toward constructing a more flexible scheduling system by focusing on more generalizable indicators among different curves. The three developed punctual indicators (morning, midday, midnight) and one durative indicator (length of day) are consistent with the concept of 'parts of the day' in common senses, which is influenced by the orbital motion of the Earth around the Sun (Clemence 1959). Using CDR data from Tallinn (Estonia), Harbin City (China), and Paris (France), we demonstrate how these metrics enable us to compare differences in time use in urban space within and between cities. A key consideration in the development of such metrics from CDR data is the selection of the appropriate level of geographical and temporal aggregation.

2. The conceptual framework

The standardization of time and time measurement of the nineteenth century was driven by the needs of industrialization and transportation networks rather than local activities or practices (Pred 1981). This history of centralizing time inspired this research to consider how we might 'socialize time' to better reflect the everyday and localized practices of society. In this effort, we build upon earlier work, such as time geography, which analyses spatial and temporal processes via an integrated framework for measuring events and activities in geographical space and is characterized via capability constraints, coupling constraints, and authority constraints (Hägerstrand 1970). We also draw upon the idea of 'social time' as defined by Sorokin and Merton (1937), which focuses on 'the rhythm and pulse of society' and stands in contrast to other conceptualizations of social time as the moments used for socializing or social interaction (Cipriani 2013). Thus, this paper defines social time as the particular structuring of the temporal dimension of society built via the aggregation of human activities and social behaviors that can be captured in various data collection systems. A key part of this conceptualization is that social time varies across scale and space and is therefore especially useful to highlight unique and localized patterns of behaviors. Conceptually, we recognize that social time varies across and within cities, e.g., Madrid's practice of siestas differs from Northern European cities and most cities have early morning markets and nightlife districts, and with CDR data it is now possible to empirically measure and visualize these differences.

Measurement of the time usage of cities can be thought of as 'demand' emerging from the behavior of people (activities of individuals) and 'supply' derived from the functioning of institutions (e.g., opening times or timetables). Both sides are co-productive of time usage with 'demand' influencing the times and places of the availability of services and the 'supply' also guiding the consumption habits and spatiotemporal behavior of the society (Bromley *et al*. 2003, Kwan and Neutens 2014). Thus, measurement of the time usage of urban space might be approached from the side of demand or supply as long as its counterpart is not ignored.

When the temporal rhythm of the human activities is measured, it is usually done via time use surveys (time observatory) or travel surveys. Such surveys are often only related to certain activities and to a limited time period. In contrast, several information and communication technologies-based big data sources create a possibility for monitoring the temporal pattern of a wide range of human activities during longer periods of time (Järv *et al*. 2014). One of the most useful sources for this approach are MNO's log files: CDR and data detail record (DDR) (Ahas et al 2008, Calabrese *et al*. 2011, Yuan *et al*. 2012). These are used extensively in research, not only because they capture a record of social interaction or activity but also because they usually record a large amount of people in a wide territory. It is difficult to capture an accurate picture of the time usage for a single person from CDR and DDR data – mobile phones are generally used irregularly – resulting in very few observations or calls for certain individuals. At a city level, however, CDR data can be temporally and spatially aggregated and thus provide a good overview of the human activities at the societal level.

In this way, aggregated CDR data can be used to derive the diurnal call activity curves of mobile network cells. At the macro-temporal level, their diurnal curves are relatively similar for most of human activity indicators: people are more active in the daytime and less active at night (Roenneberg *et al*. 2003). Geographically, however, divergent patterns emerged. For example, suburban 'sleeping areas' and jobs and business centers have temporally inverse patterns, with the latter containing most activities during the working hours and the former experiencing activities primarily outside of the work day (Meijers 2007). Services, recreation areas, and scenes for cultural or sport events can be distinguished in the same way, albeit with different temporal concentrations of activity. Thus, comparison of time usage curves enables us to compare the social times of areas and people.

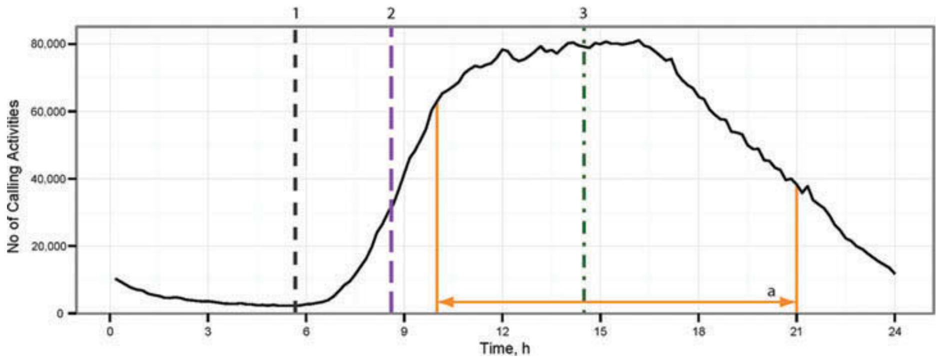

Figure 1. Diurnal call activity curve which presents the amount of call activities in every10-minute time period (all together 144 periods per 24 h) in Tallinn, Estonia. Presented aggregated data for all network cells in city: 1 – night minimum; 2 – morning, 3 – midday, a – duration of day.

Figure 1 presents 'a classic' CDR data-based aggregated time usage curve for the city of Tallinn, Estonia, in which the number of call activities of the mobile phone users in the city are aggregated in 10-min periods, 144 periods per day. Diurnal distribution of call activities can be also standardized. Figure 1 also provides an example of the three fixed time periods – night minimum (line 1), morning (line 2), and midday (line 3), as well as the duration of the active day (the shaded area labeled 'a') that we use in this analysis.

Such 'diurnal call activity curves' based on CDR data do, however, have limitations (Yuan and Raubal 2012). Comparing graphs visually highlights the differences between the time usage of certain areas or groups, but statistical testing is complicated as it is challenging to find reasonable dependent values from these diurnal call activity curves.

The literature on temporal phenomena focuses on linear and cyclic processes (Feldman and Hornik 1981). The linear processes include, above all, long-term variations in demography, land-use change, migration, etc. Cyclic processes are mainly short-term, usually with a 24-hour, weekly, or seasonal cycle and include such things as commuting, tourism, seasonal employment, and agriculture in the case of climate zones, etc. (Panda *et al*. 2002, Roenneberg *et al*. 2003, Silm and Ahas 2010). Cyclic processes are rhythmic, i.e., the events within the phenomena occur with a certain regularity with periods and amplitudes, albeit with range of variance. The rhythm of the processes is thus determined by the repetition of certain events that can be fixed and measured in time, such as the practice within climatology to fix the start of winter to the date of the formation of permanent snow cover (Jaagus and Ahas 2000). The duration of events can also be measured, such as the definition of winter measured by the period of time with permanent snow cover. Building upon this standard practice, one can also study the rhythmic processes of time usage within and between cities through the identification of measureable events of urban life. The goal is to both confirm generalized patterns shared across urban areas and identify how the social time geographies of each city differ.

Cities have different functionalities, which are temporally used in very diverse manners. Employing temporal urban activities as an indicator of vitality has long been deemed important in the works discussing the urban monofunctionality (Filion 2000). Monofunctional areas are attractive to people in a short period during a day and stand in contrast to multifunctional areas, which are active for longer stretches of time. The so-called' 24-hour city' stresses the importance of 'mix use' and night-time entertainment as a tactic to extend the usage time of a city and has become a key concept for the temporal dimension of a city (Bianchini 1995). Thus, it comes as no surprise that the 24-hour city concept has been utilized in urban planning and development strategies to help invigorate the economy of a city (Lovatt and O'Connor 1995). In addition to the circadian cycle, other important urban rhythms include weekly (7-day) and seasonal (12-month) cycles.

In travel behavior research, activities are divided into home, work, and leisure, with each activity characterized by a specific temporal and spatial pattern correlating with land use (Forsyth *et al*. 2007). Cities not only contain monofunctional areas such as suburban 'sleeping' districts and industrial 'work' areas but also include many multifunctional areas as well with various activities entwined in the same buildings and/or districts. For example, some people may work at home while conversely many household activities are taken care of at work during the day. In short, urban environments are crossed by various trends in the use of space and time (Handy *et al*. 2005) or, in the parlance of this research, cities possess unique and localized geographies of social time.

3. Data and methods

3.1. Data

To test our methodology for mapping differences in social time in the case study cities – Harbin in China, Paris in France, Tallinn in Estonia – we used CDR data provided by the major mobile operators in Estonia (EMT), China (China Mobile) and France (Orange). Data for the Chinese cities cover 9 days between 21 and 29 July 2007, while French and Estonian datasets are drawn from 22 to 29 September 2007 to avoid potential behavior inconsistency during the European summer vacation period. Sunrise in Harbin City was at 4:03 am on 21 July 2007; in Tallinn 07:04 am and in Paris 07:36 am on 22 September 2007. We studied working days (Monday, Tuesday, Wednesday, Thursday, Friday), weekend days are different but we do not compare them here. The Estonian data only includes outgoing call activities (call, text message, data communication, services) and averages 5.6 activities per person per day. Data from China and France includes all incoming and outgoing call activities with an average of 5.5 activities per person per day in China and 8.13 in France. In total, the database includes 2.4 million users in Harbin province in China, 1.7 million in France, and 0.3 million in Estonia.

Although it is beyond the scope of this paper to provide a definitive explanation, it is clear that social time of these cities is influenced by culture, traditions, the political system, and the urban context. For instance, the three cities in this study differ to lesser or greater extents along these dimensions, all of which may contribute to the observed time usage. Harbin City is an important industrial area of the rapidly developing China. It has 5.3 million residents (10 million if the surrounding suburbs are included) with most living in high-rise apartment buildings serviced by centrally planned urban structures and transportation. Paris is a classic western European city with 2.3 million residents in the core city and more than 10 million in the entire urban region and is deeply embedded within the time usage traditions of French culture and society. Employment reflects the so-called postmodern shift, as the number of services, education, and creative job positions – associated with more flexible temporality – are increasing relative to industrial employment. Tallinn, Estonia, has 400,000 people living in the core with another 200,000 in the surrounding hinterland and is the center of a rapidly developing Eastern European country. Tallinn is still transitioning from a Soviet centrally planned society to a market-based economy, with accompanying changes to employment structure and daily activities as the traditions and rules in the society are being reinvented. This variety of urban contexts helps to ensure that our methodology for analyzing social time is relevant across a wide range of spatiotemporal patterns and differences.

3.2. Data aggregation

In order to establish a uniform time unit for comparison, we aggregated the data of call activities into 10-minute periods from 00:00 to 24:00. We normalized the results relative to the average distribution of call activity in a day using a metric from zero to one with zero indicating that no calls were made and one indicating that the maximum amount of activity took place. The resulting dataset provides a series of 144 points for call activity within each city with a 10 minute temporal resolution (Figure 1). In order to stabilize the noise and extract indicators, we also smoothed the curves based on local polynomial regression fitting (LOESS), which utilizes the local subset points to create a fitted value (smoothing parameters: LOESS (count ~ time, span = 1/3)). The major advantage of LOESS is that it does not require a specific function to fit a model to the entire dataset; therefore, it is ideal for modeling complex phenomena where no theoretical models have been established (Ripley 2004). For this study, it is challenging to fit a specific model into the varying call activity curves; hence, LOESS is selected as a smoothing technique.

3.3. Key metrics of time

In order to study the time use differences (or social time) in cities, we use four indicators: (1) night minimum; (2) start of the morning; (3) midday; and (4) duration of active day (Figure 1, B). Note, all times referenced in this article are local standard time for the specific city mentioned.

The 'night minimum' indicator is defined as the moment with the lowest number of calls during the 24-hour daily cycle, and represents the time point when a city is least active. In case there is more than a 10 minute period with the minimum time, the average time of these periods is calculated. (Figure 1, 1)

The 'morning' indicator corresponds to the time when call activity is increasing the fastest during the period from 3:00 to 12:00. This is the point at which the first derivative (rate of change) of the data series reaches its maximum. (Figure 1, B,2)

The 'midday' indicator is calculated as the average time of all call activities. It is found by summarizing all values of 144 ten-minute time units of the diurnal cycle and dividing by 144 (Equation (1)). This average time indicates the 'mean' of temporal distribution of activities in a city and is easy to determine. (Figure 1, B,3)

$$t_x = \frac{\sum ca}{N} \tag{1}$$

where: $t_x = midday$
$\sum ca$ = sum of calling activities time
N = Number of calling activities

The ‘duration of active day’ indicator is calculated as the period in which 80% of all call activities are made. Starting at midnight, the active day begins when 10% of activity has occurred and ends when 90% of activity has occurred. For example, if 10% of cumulative call activities is reached at 8:20 am and 90% at 6:33 pm, the duration of active day is 10:13. (Figure 1, a)

3.4. Spatial analyses

Analyzing the spatial differences of social time in cities is most easily achieved by mapping the timing variables outlined above. Such an interpolated map for the midday marker of social time in Tallinn is shown in Figure 2; however, this relatively simple approach can be characterized as ‘fuzzy’ and does not provide a good overview of the distribution of social time in a city. The lack of regularities in Figure 2 is caused by the high density and diversity of mobile network cells. There are regular antennae located in the urban space, there are antennae located in important buildings (such as shopping centers, rail stations, etc.), and there are antennae that are located in the vicinity of larger buildings, markets, and transportation centers. This means that the coverage of network cells is not universal or uniform, as the distribution of the cells matches the functional diversity of the urban space and buildings, and thus the visualization in Figure 2 provides poor guidance in understanding the spatial patterns of temporal activities.

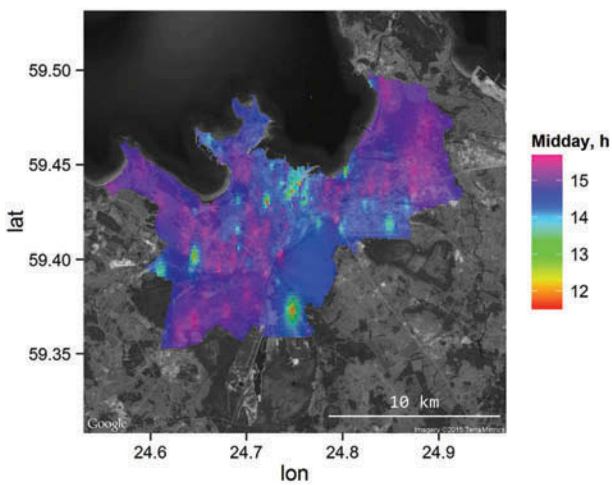

Figure 2. The midday period on weekends expressed in standard time in mobile network cells in Tallinn. Spatially interpolated by the inverse distance weighting (IDW) method.

Thus, after several tests, we addressed the issues shown in Figure 2 by mapping the beginning times for our social time indicators through the use of quartiles. This method is suitable for studying the temporal dynamics of the urban space as we can compare the course of a certain temporal event – for example, midday – by tracking the distribution of network cells within each quartile. The maps are composed from network cells with first quartile = very early; second quartile = early; third quartile = late; fourth quartile = very late beginning time. This level of temporal aggregation reduces the effect of varying distribution of antennae in space. We presume that activities of more similar temporal distribution in a city are represented in the quartiles.

To visualize the geographical differences in social time in urban space, we first determined the geographical center of the city by calculating the geometrical center of the mobile network antennae used in the study. From this geographical center, we mapped the ellipses for the network cells associated with each quartile in meters. Even though the geometrical center of the mobile network antennae does not precisely match the center of the city, it is clear for finding a central point in the system and readily shows spatial differences between quartiles. Moreover, many modern urban geographical approaches recognize the difficulty in determining the center of a city, particularly in the case of polycentric urban areas, and city centers have many definitions and methods (Meijers 2007).

4. Analysis

4.1. Duration of day

Comparison of the duration of days of the cities shows that the active day is the shortest in Paris, (10:25) and longest in Harbin (11:41) (Table 1). The values of the standard deviations and medians in Table 2 show that variation between mobile network cells within the cities is not very high. However, the large amplitude of the minimum and maximum values (minimum in Paris 3 h, maximum in Tallinn 22:40) also shows that the

Table 1. Descriptive statistics for day length for all network cells with the cities.

City	Mean	Median	Min	Max	SD
HARBIN	11:41	11:50	07:10	14:00	00:52
PARIS	10:25	10:40	03:00	18:30	01:07
TALLINN	10:30	10:40	07:00	22:40	01:31

Table 2. Statistics describing arrival of midnight.

City	Mean	SD
HAR	03:24	01:05
PAR	05:16	00:48
TAL	05:11	00:48

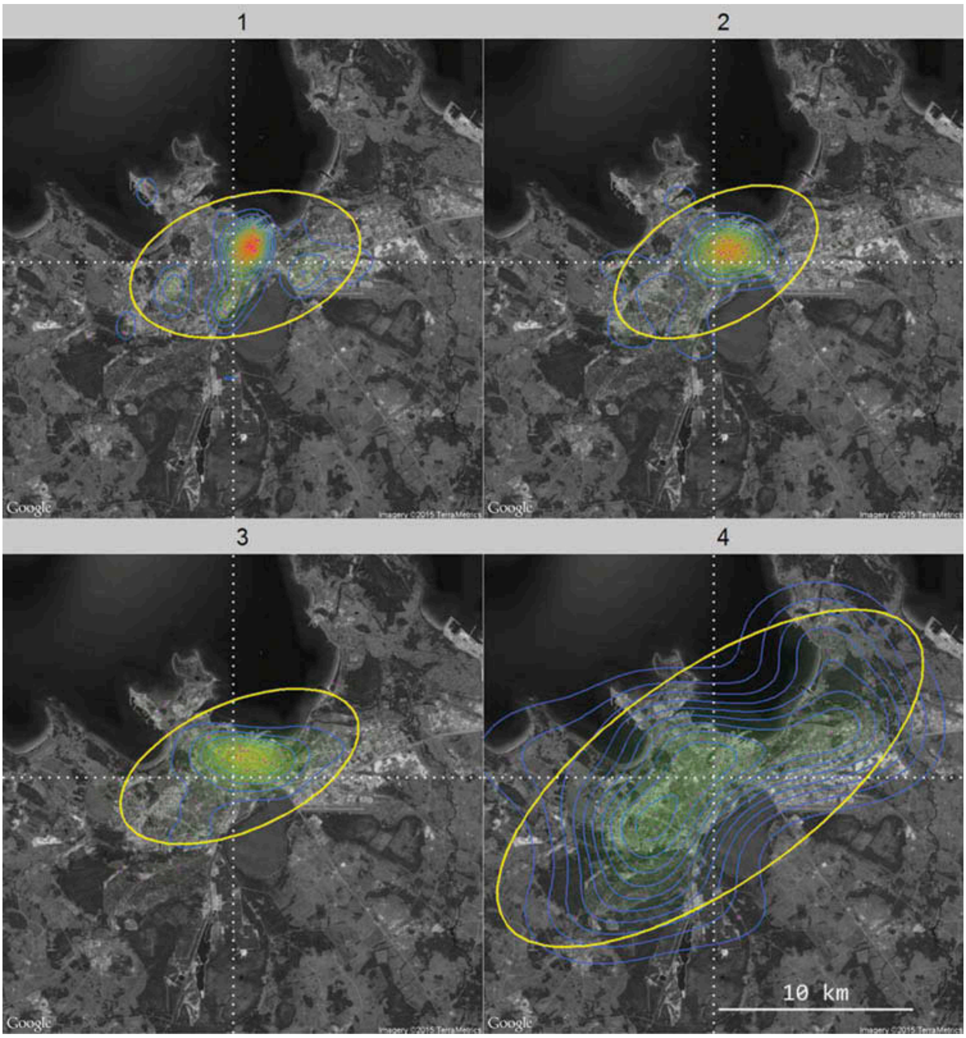

Figure 3. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of day length in Tallinn.

network cells of certain locations or buildings may be in limited use temporally or spatially.

This difference between day lengths of network cells in the first and fourth quartiles also varies between cities: the value for Harbin is 2:15, Paris 2:21, and Tallinn 3:17. With the largest difference in day length between quartiles, spatial differences in duration of day are the greatest in Tallinn: the active day is the shortest (quartiles 1, 2, 3) in a narrow area in the city center on a working day; while network cells associated with the longest day (quartile 4) are dispersed over Tallinn (Figure 3).

In Harbin City (Figure 4), the geographic differences of the duration of the active day quartiles are much smaller. The city center differentiates with the shortest day (quartiles 1,

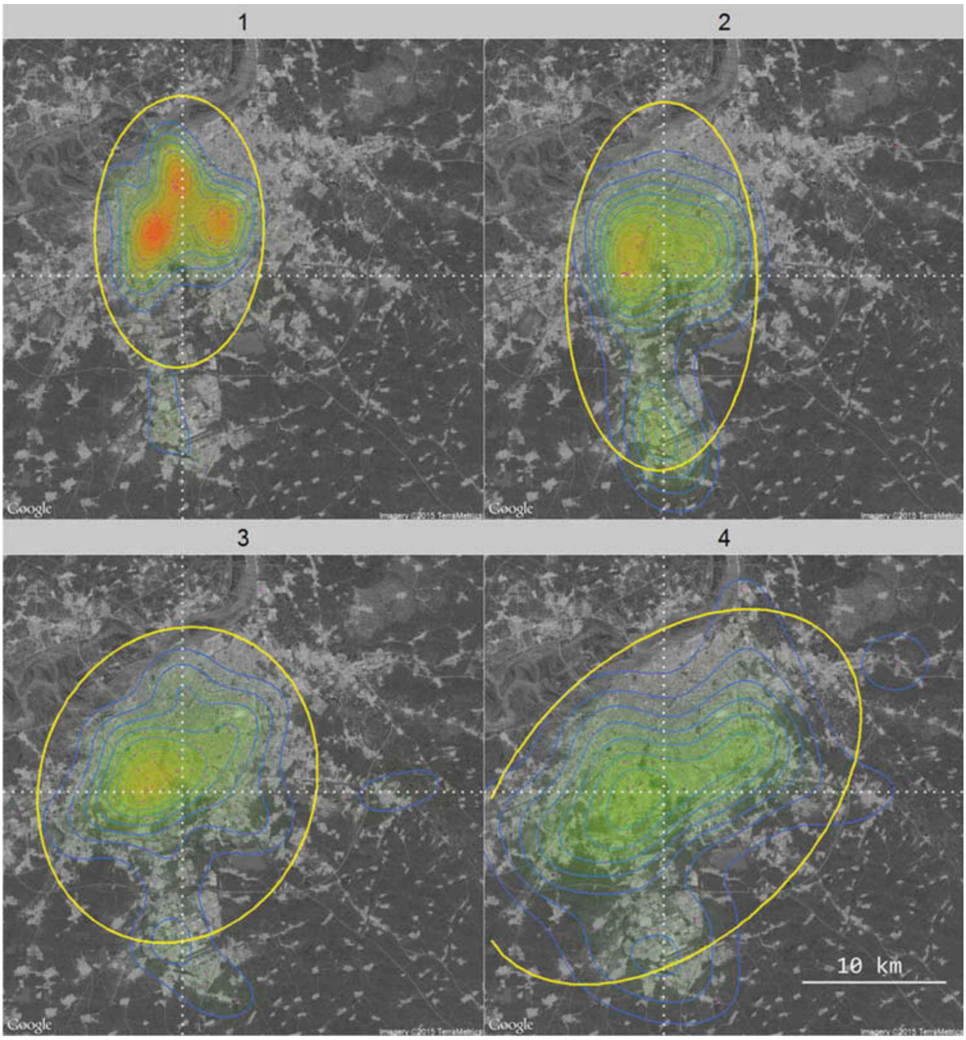

Figure 4. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of day length in Harbin.

2, 3) and the antennae with the longest day (quartile 4) are more diffused over the city. In Paris (Figure 5), the differences in the geographical distribution of the duration of active day quartiles are greater. On a working day, the shorter day (quartile 1, 2) is focused in the center of the city, the antennae with a longer day are, however, diffused all over the city (quartiles 3, 4).

This difference between cities is illustrated in Figure 6, which presents the duration of the active day by quartiles. The shaded area around the lines marks the general variation in day lengths within the quartiles. Quartile 1 (very early) varies most. These temporal differences between quartiles also have geographies, such as a narrow concentration in the center of Tallinn as shown in Figure 3.

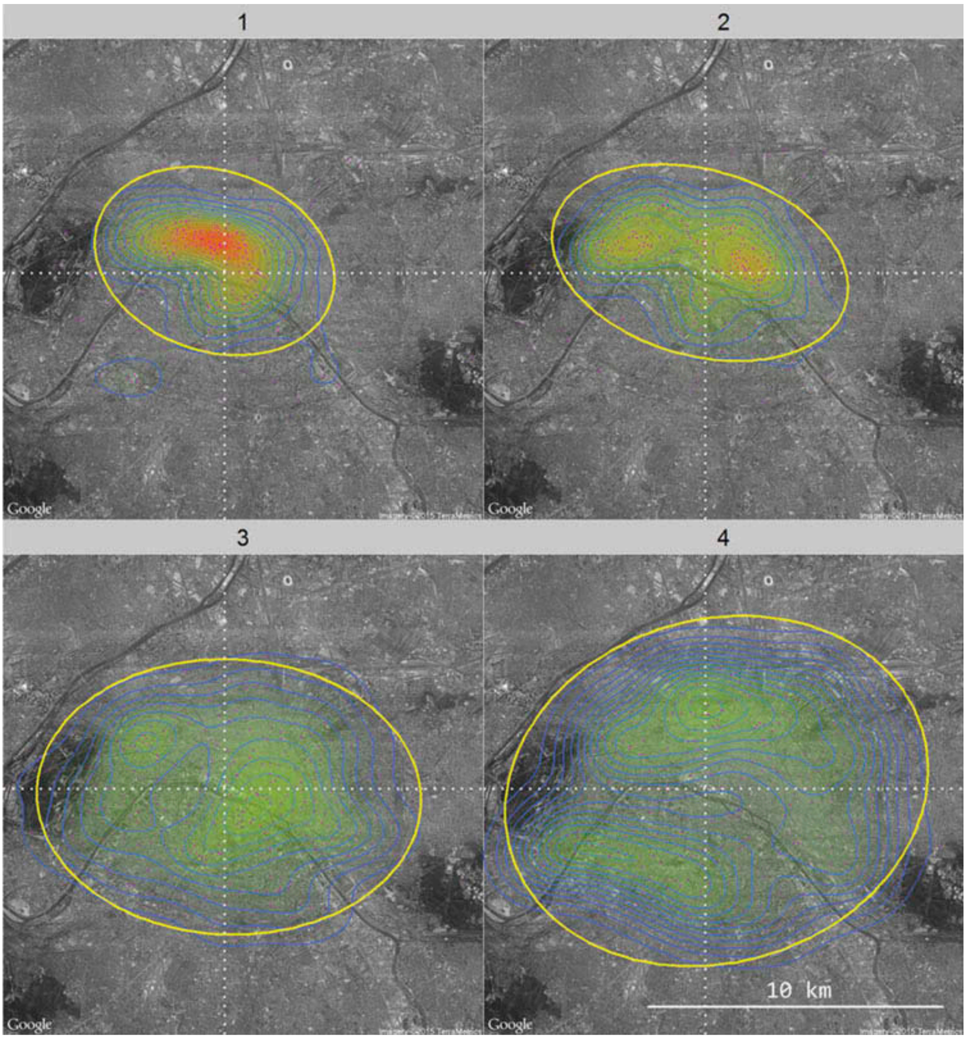

Figure 5. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of day length in Paris.

4.2. Night minimum

The night minimum of phone usage shows the least active moment of a city. In Harbin, the minimum of the city arrives earliest, at 3:24 am on working days. In Paris and Tallinn, the night minimum only arrives in early morning, i.e., at 5:16 am in Paris and 5:11 am in Tallinn (Table 2). Variation of the night minimum within the cities is moderate at around 1 h SD.

The difference between the averages for network cells in the first and fourth quartiles in Harbin is 2:34, in Paris 1:46, and in Tallinn 1:50. In Harbin, the beginning of the first quartile of the night minimum is diffused over the city, the second quartile is more concentrated in the city center, the third and fourth quartiles, however, are concentrated in a very narrow city center area (Figure 7).

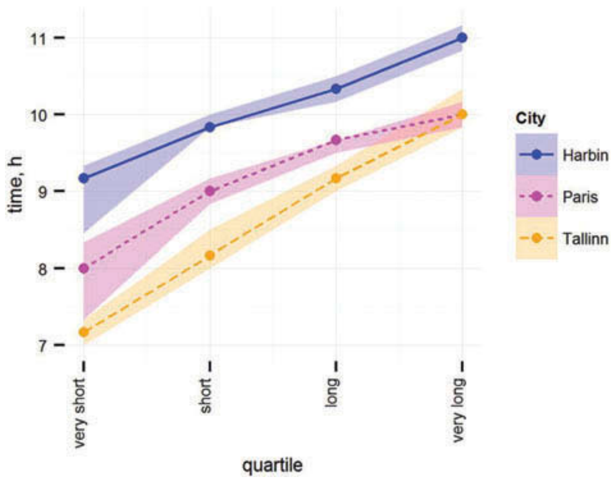

Figure 6. Distribution of day length in the studied cities by quartiles.

In Paris, the geographical distribution of midnight is quite similar to Harbin: the arrival of the first quartile is diffused over the city and the fourth quartile is concentrated in the city center. The geography of the arrival of the night minimum in Tallinn follows a pattern similar to that of Harbin and Paris. Graphic presentation (Figure 8) of variation of the quartiles shows that in Harbin, the variation of the night minimum with the fourth quartile network cells is very large.

4.3. The morning

The morning starts with a rapid rise of activities in all cities. In Harbin, the morning starts at 5:00 am, in Paris at 6:20 am, in Tallinn at 5:57 am, and there are remarkable differences in this indicator of social time between cities and quartiles within cities (Figure 9, Table 3).

For example, the geography of the start of the morning in Paris varies with network cells within the first quartile dispersed over the city (Figure 10), the antennae of the second quartile are concentrated north from the city Centre, the locations of a later start of the morning, i.e., the third and fourth quartiles, are in the areas north and northeast from the city center.

The morning in Harbin also starts with great spatial variability; the first quartile is diffused over the city, the second northeast from the city center, the third west from the city center, and the later start of the morning (fourth quartile) is concentrated in the north and northeast from the center. The start of the morning in Tallinn is distributed over the city in the very early and early quartiles and concentrated in a narrow area in the city center in the late and very late quartiles.

4.4. The midday

At midday, the average time of all call activities – is quite stable in all cities (Figure 11). Harbin is the city with the earliest midday (at 2:03 pm on a working day), followed by Tallinn (at 2:33 pm), with the latest midday in Paris (at 3:38 pm) (Table 4). The differences between the quartiles of the arrival of midday are greatest in Tallinn. In

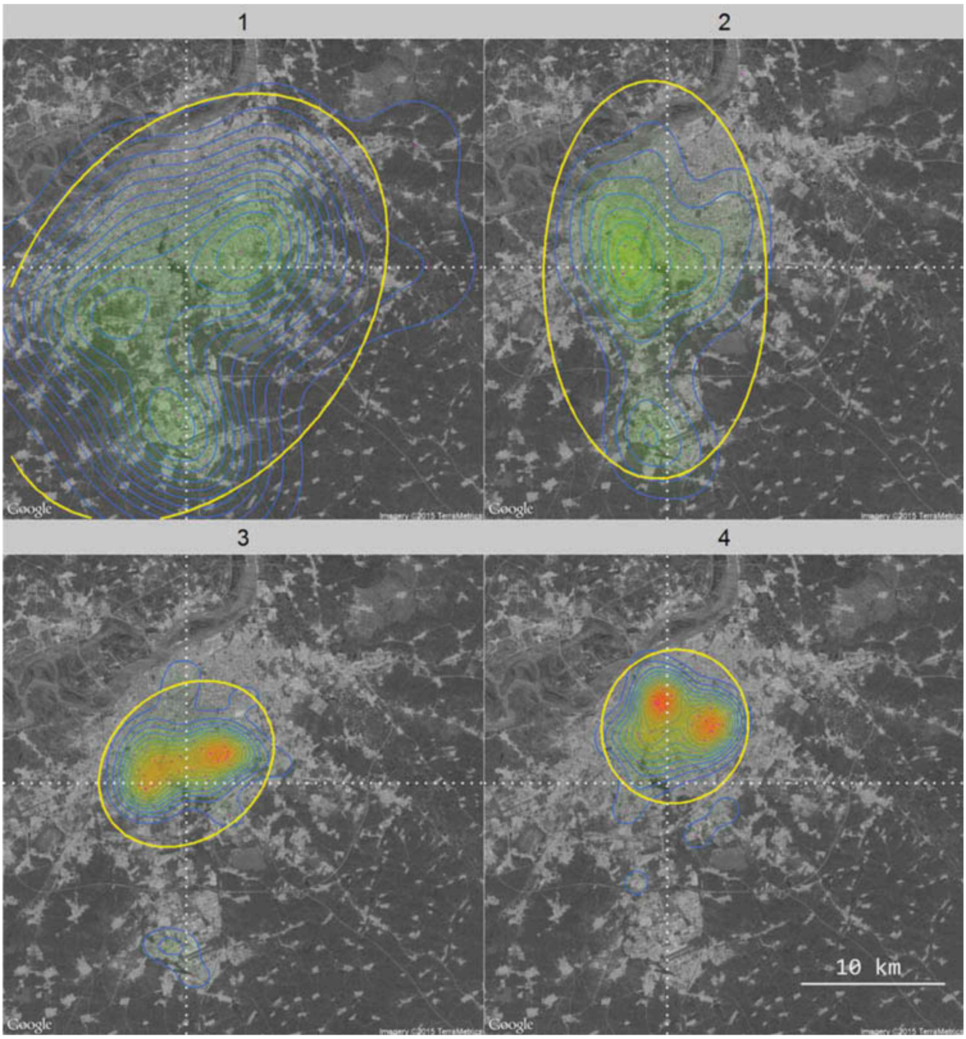

Figure 7. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of night minimum in Harbin.

Harbin, the same indicator remains within 1 h on both types of day and in Paris within 1.5 h on both types of day. Intra-quartile variation of midday is the greatest in the first quartile in Paris and Tallinn (Figure 11).

Geographically, arrival of midday in Harbin is distributed quite evenly over the city with concentration in the city center. Only in the case of the very late (fourth) quartile are the antennae primarily concentrated in the city center. In Paris, the distribution of midday quartiles is more dynamic. The first and second quartiles of a working day are concentrated in the north of the city center; the third and fourth are more dispersed east from the city center. In Tallinn, the first, second, and third quartiles on a working day are quite narrowly concentrated in the city center, the fourth is dispersed over the city (Figure 12).

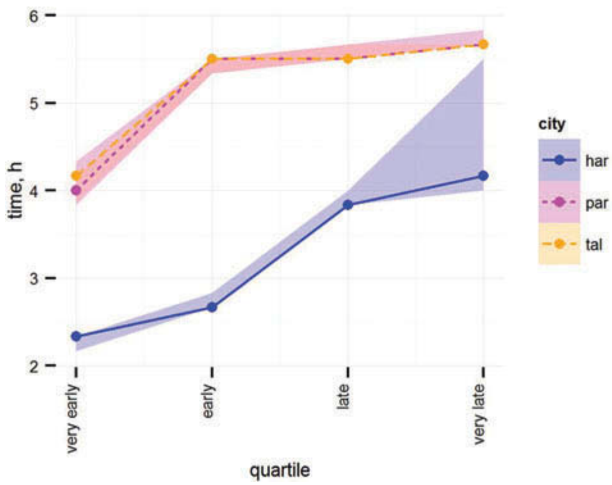

Figure 8. Distribution of the night minimum in the studied cities by quartiles.

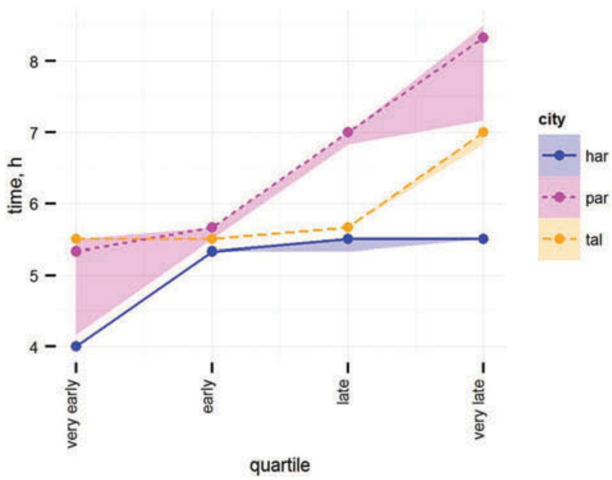

Figure 9. Distribution of the start of the morning in the studied cities by quartiles.

Table 3. Statistics describing morning.

City	Mean	SD
HAR	05:00	00:43
PAR	06:20	01:17
TAL	05:57	00:46

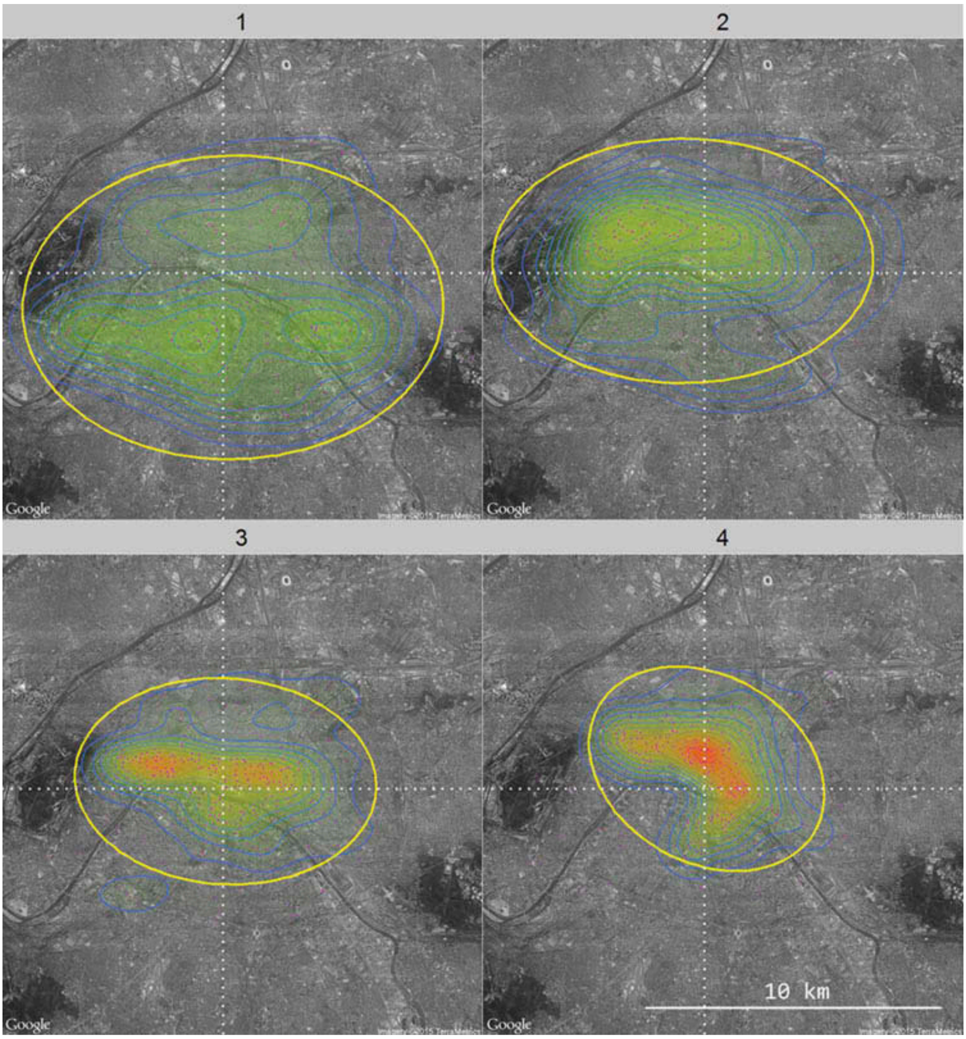

Figure 10. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of start of the morning in Paris.

5. Discussion

5.1. Evaluating the indicators

Mobile phone data is not a novel source for time use studies, although earlier researchers have mainly used these data for studies examining diurnal or weekly time use curves. The methodology developed in this paper takes a different approach, and measures the timing (particularly the start times) and duration of certain events, replicating the insights offered by climatology and seasonality studies (Jaagus and Ahas 2000) in an effort to construct localized metrics of social time. In developing the methodology behind identifying useful indicators within the diurnal cycle, we explored a range of timing metrics. Based on these tests, we selected three indicators of timing (night minimum, morning, midday) and one period (duration of day) as the best-suited ones for studying social time and, by extension,

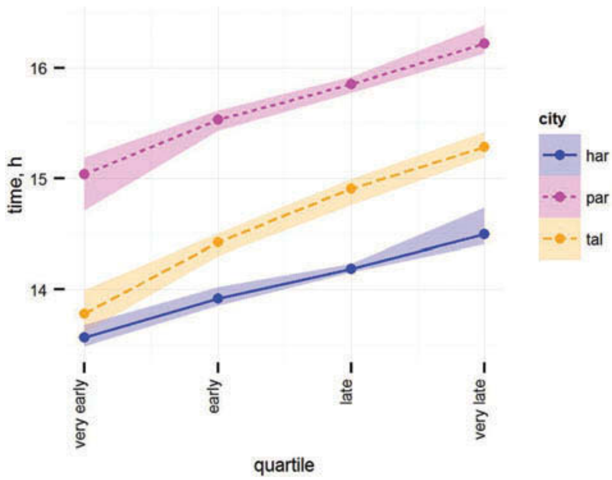

Figure 11. Distribution of midday in the studied cities by quartiles.

Table 4. Statistics describing midday.

City	Mean	SD
HAR	14:03	00:25
PAR	15:38	00:40
TAL	14:33	00:50

the temporal and spatiotemporal differences between and within cities. All of these metrics based on CDR data provide replicable means for measuring temporal variation in the activities in a city.

In order to test our methodology, we used CDR data from Harbin, Paris, and Tallinn to analyze differences within and between cities. The analysis showed that the most significant problem is the shortage of data in certain network cells. Smaller network cells and sparser population mean that there are simply too few call activities and it is not possible to calculate the indicators adequately at this scale. While this was not a particularly big problem in Harbin and Paris (both with sizeable populations), it emerged as a significant issue within the much more sparsely populated Tallinn. This was especially a problem in the determination of night minimum, as some sparsely populated areas could be without calls during multiple 10 minute time periods during the daylight hours of the day. This situation can be addressed by aggregating data in network cells, extending the time period used or, as we did in this case, setting a temporal frame from 01:00 to 08:00 for the determination of midnight.

In addition, there are specific issues associated with each indicator of social time. Another concern with the night minimum is that even with the temporal frame, the period with a minimal number of calls may be long or spread across several minima. In these cases, the central point in time for multiple minimum periods is calculated. An issue for the morning phase is the occurrence of periods with a similar curve in other parts of the day. For example, in Paris and Harbin, similar derivate values were detected in some cases during fast-growing call activity arising right after the lunch break. This again was

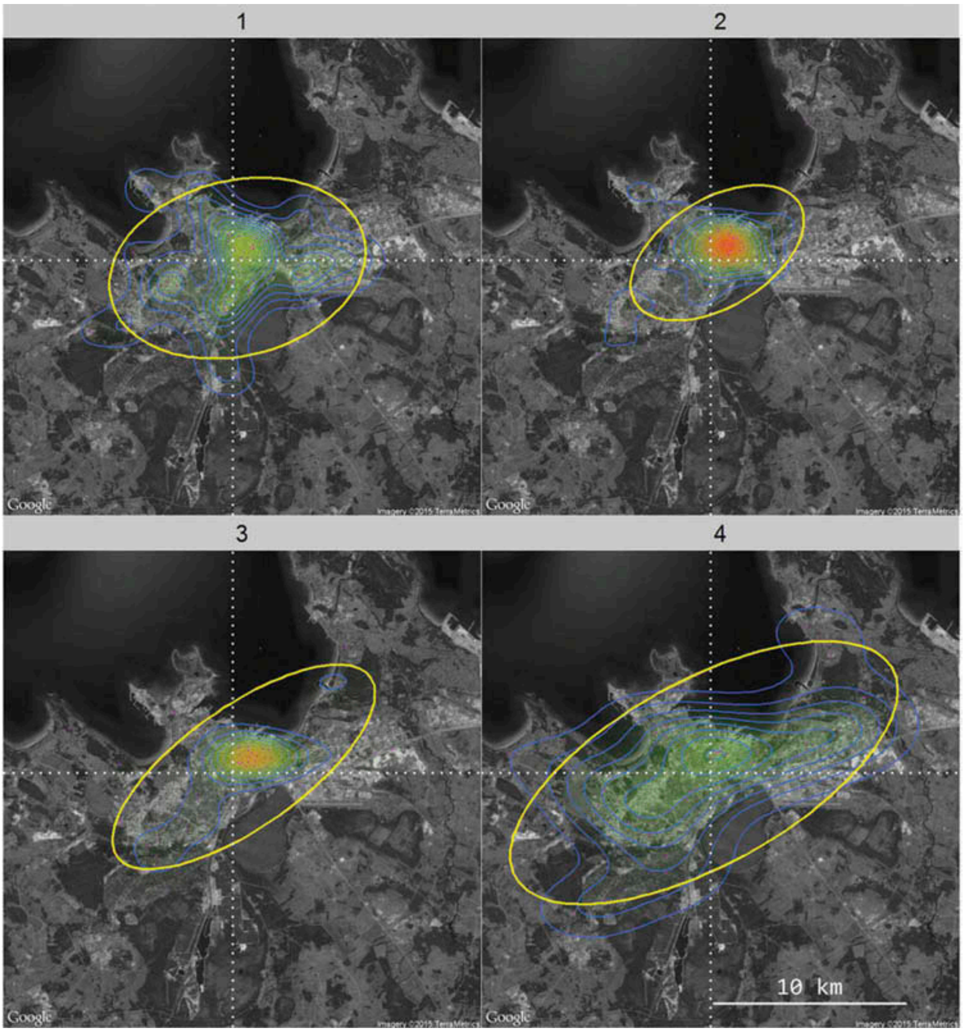

Figure 12. Relative density and standard deviational ellipse (1 SD) of mobile antennas by quartiles of midday in Tallinn.

addressed by defining a time frame for the start of morning: in this case, 3:00 to 12:00. The calculation of time of the midday and duration of the active day was methodologically easiest and thus serves as a particularly good metric for comparison of inter- and intracity activities. But the indication and interpretation of these indicators would benefit from more theoretical explanation as this line of research moves forward.

5.2. Spatial analyses

Spatial analysis of the four indicators revealed that due to the very high variability in values of individual network cells, it is simply not possible to discover timing regularities in cities with a simple mapping (see Figure 2). The social time in urban space revealed by

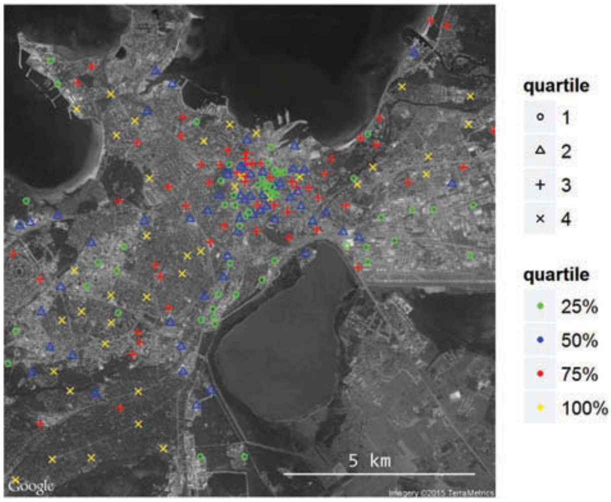

Figure 13. Day length in Tallinn, values of four quartiles of Figure 3 presented on one map.

these metrics can be very spotty, there may be a 'very late' antenna in the vicinity of a 'very early' antenna, and the values are too different for discovering any spatial regularity. To solve this problem, we developed a methodology for dividing time events into four quartiles, which helps smooth the visualization shown in Figure 3. However, even plotting such quartile data on a single map – such as the one in Figure 13 – results in a representation that is difficult for interpretation and analysis of the spatial differences in temporal activity.

We also tested mapping with terciles and quintiles, but with the current data the most effective division was quartiles. The cartographic analysis of the quartiles (very early, early, late, and very late) of beginning times of events provides a good overview of the areas with early and late timings in the city and shows that the timing of human activities in the city has different temporal layers. Such deconstruction of data layers into quartiles can be a useful tool for various spatial analyses. This method also has limitations, as we can only see general patterns of distribution of social time and may miss interesting locations and events in the city. But this is a problem of selecting the appropriate method for achieving specific research objectives.

5.3. Differences between cities

The results of our analysis highlight the substantial differences in the social times of the three case study cities. The duration and timing of the CDR-derived markers vary across different cities and geographical areas in cities. The three temporal indicators studied (midnight, morning, midday) show that the day in Harbin arrives and ends later and is therefore longer than in Europe. Some of these differences can certainly be tied to geographical location, seasonality, and issues with standard time. However, it is also reasonable to suggest that cultural differences (including economic, governance, traditions, etc.) also contribute to variation in time use. The earlier start and longer days in Harbin are consistent with our understandings of industrialization and fast-growing economies, while the later starts and short days correspond to the postindustrial societies, such as France and Estonia. Thus, the indicators developed in this study provide an intriguing first step for future research on the causal relationships between social time and spatial patterns of activity, land use, and other factors. Another important factor likely influencing the results of this study includes the size and density of the urban areas. In Harbin, the study area covers the whole city, including high-rise apartment buildings and densely populated suburbs. The area included in the study in Tallinn covers the city center and the suburbs of lower buildings and sparser population surrounding the city. In contrast, the study area in Paris only covers the core area of the city. In Paris, the people leave the area after work and the day seems shorter to us and the timing different. Such differences highlight some of the challenges of using CDR data from multiple countries and demonstrate the aspects that must be taken into account when comparing different data and places.

5.4. Time use differences within cities

Our analysis shows that social time differences between the city center and the suburbs are specific to each city. The very early (first) quartile and the very late (fourth) quartile have opposite geographic distributions in city centers and suburbs. Functional difference between suburban–urban structures and activities is a well-studied aspect of urban life, and our results confirm that methods based on mobile phone data are useful for detecting and measuring those differences. In addition to the quartile maps (Figures 3, 6 and 10), it is also interesting to present the densities of activities graphically as the distance from the central point of the city. Figure 14 shows the geographical distributions of morning start in the studied cities. The peak for the earliest (first) quartile of the start of morning is the furthest away from the city center in Harbin, and the following quartiles are each closer to the city center with relatively different values for the kurtosis, skewness, and shape of the distributions. In Paris and Tallinn, the distribution of social time in space is similar, although the spatial differences in Paris are simply smaller (a product of the size of the study area) and, in Tallinn, the latest quartile is much more highly concentrated in the city center. Presumably, this means earlier waking of the suburban sleeping areas further away from the center and later arrival of people to the workplaces in the city center. Of course, in interpreting this result, we must remember that the city center is defined as the geometric center of the mobile network antennae.

5.5. Sensor data and urban monitoring systems

The final goal of this paper is to prepare CDR-based indicators and methodologies for developing urban sensor systems and monitoring tools. One crucial issue in developing sensor systems for urban environments is related to communication and data management standards (Hancke *et al*. 2013). A big advantage of the Digital Mobile Network Operator data used in this study is that it has no problems with standards of communication, making data collection and management technically feasible across a range of settings. This key strength of mobile data makes it attractive for developing smart city solutions, although problems related to data access and concerns related to privacy remain (Positium 2014). A peculiarity of CDR data is that it provides a robust yet limited metric of activity; we know only location points of individuals at the moment of communication (place, time, ID). But this limitation is compensated by other features of mobile data: (1) the number of users and observation period are high; (2) data is digital

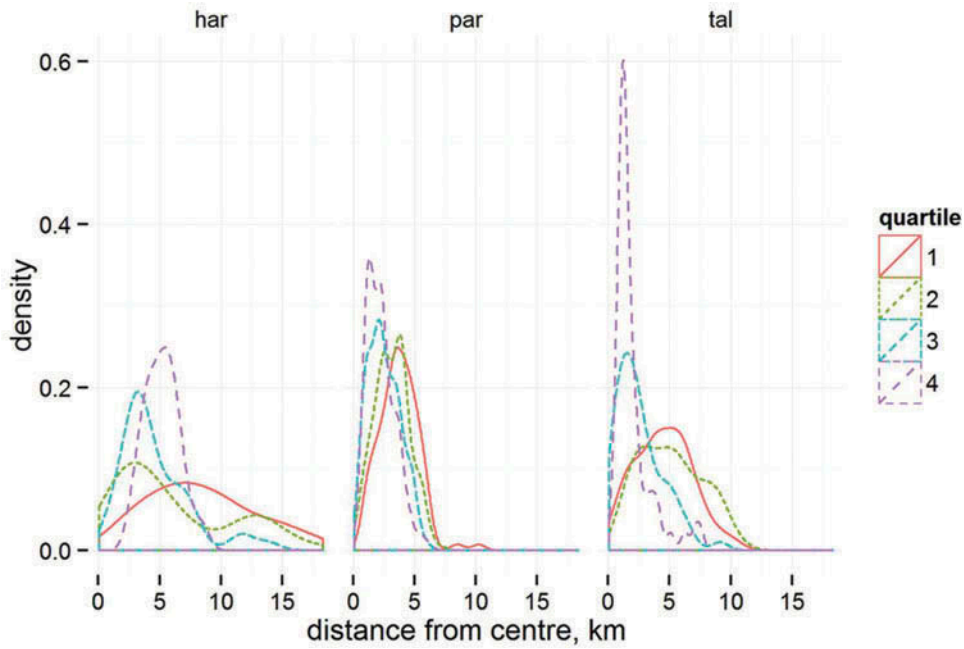

Figure 14. Distribution of the start of morning, based on quartiles, with the distance from the central point of the city.

and easy to process; and (3) the data has real-time applicability. These three aspects make mobile data an especially interesting source for sensor technologies. This robustness of mobile data can be further enriched with methods such as those developed in this study. Timing indicators derived from raw data can help to detect important spatial and temporal processes in city.

Mobile data can be used as sensor input in different levels of smart systems. ‘Robust’ location data is an important part of ‘Basic Smart’ systems and is used for sensing, monitoring, acting, and controlling functions (Debnath *et al*. 2014). This ‘robust location’ information of individuals provides a valuable base model to link to different features from mobile devices or persons carrying phones. The timing indicators developed in this paper can be used as part of ‘Advanced Smart’ systems of smart cities – for predicting, healing, and preventing functions (Debnath *et al*. 2014). For example, the social time indicators developed here can be applied to the temporal adjustment of origin–destination matrices in intelligent transportation systems and for predicting transportation demand for separate weekdays or seasons. Indicator of ‘night minimum’ can be used for measuring the actual start time (and end) of day, which can be a very different time from standard 00:00 midnight. Such ‘functional midnight’ can be used for developing behavioral taxation systems, or intelligent transportation, parking, street cleaning systems, etc. Our experience demonstrates that we can discover much higher variability of transportation parameters from analyses of long-term mobile phone time series than using short-term questionnaires or transportation census (Järv *et al*. 2014). Ubiquitous mobile data is detecting many more activities than we can grasp with traditional data collection methods.

Our methodology for quartile-based spatial deconstruction of time layers could also be applied to monitoring urban changes as part of ‘Basic Smart’ systems as well as for developing ‘Advanced Smart’ systems of smart cities. The easiest examples of such preventative measures include influencing time use in cities with automatic lighting; influencing duration of activities in urban space with dynamic parking prices and functioning times of services. Such ‘Advanced Smart’ systems can be utilized in cooperation of urban sensor systems, the internet of things and ‘smart city infrastructure’ (Atzori *et al.* 2012).

6. Conclusions

Our objective in this research was to develop a methodology for measuring time use patterns of urban life (or social time), using mobile communication datasets, which could be useful in urban monitoring tools and smart city solutions. Using analogs from climatology, we developed three temporal indicators (midnight, morning, midday) and one time period (length of day) for measuring diurnal rhythms in a city. We also developed a quartile-based approach for spatial analyses of timing in urban space and tested these methods using datasets from Harbin, Paris, and Tallinn.

Our results highlight that there are differences in social time patterns across these cities, each with its own characteristics and culture. In Harbin, activity starts earlier and active days are longer than those in either European cities in which activities in the city centers last for longer periods of time. Our indicators were also very good in identifying time use differences within each city. There were clear differences in time use patterns between the city center and suburban areas, which is also an area of interest for further study on the causality between urban functions and human activities.

This paper offers an initial step in the use of CDR data in measuring and understanding the spatiotemporal variation between and within cities in the age of instant access. This particular case study provides an initial thought for developing different urban monitoring tools and smart city solutions that rely upon fast data feeds and the automatic detection of urban change. Thus, it is in this context that the use of CDR data and our indicators are the most relevant: especially in the application of ‘Basic Smart’ and ‘Advanced Smart’ systems governing urban space. These activity-based indicators provide new sources for determining more flexible time standards in modern societies. The methodology in this study can be extended in several aspects. For instance, besides the four general indicators in this research, it is feasible to define more detailed indicators based on the specific aggregated curves, e.g., we can define lunch break as the local minimum point around noon, which corresponds to the time when the activity is the lowest [10:00–14:00]. It is also valuable to further validate the robustness of the proposed methods with richer data and more complex scenarios.

Acknowledgments

We wish to thank the anonymous reviewers for their helpful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by Estonian Research Council Institutional Research Grant [IUT2-17]; EU Regional Development Foundation projects TERIKVANT [3.2.0802.11–0043]; Real-time Location-based Big Data Algorithms [3.2.1201.13-0009]; Swiss National Science Foundation [Project 205121_141284]; United States Department of Transportation Eisenhower Fellowship Program; and European Regional Development Fund through the Software Technology and Applications Competence Centre.

ORCID

Y. Liu http://orcid.org/0000-0002-0016-2902

References

Ahas, R., *et al*., 2007. Seasonal tourism spaces in Estonia: case study with mobile positioning data. *Tourism Management*, 28 (3), 898–910. doi:10.1016/j.tourman.2006.05.010

Ahas, R., *et al*., 2008. Evaluating passive mobile positioning data for tourism surveys: an Estonian case study. *Tourism Management*, 29 (3), 469–486. doi:10.1016/j.tourman.2007.05.014

Atzori, L., *et al*., 2012. The social internet of things (SioT) – when social networks meet the internet of things: concept, architecture and network characterization. *Computer Networks*, 56, 3594–3608. doi:10.1016/j.comnet.2012.07.010

Batty, M., *et al*., 2012. Smart cities of the future. *The European Physical Journal Special Topics*, 214, 481–518. doi:10.1140/epjst/e2012-01703-3

Bianchini, F., 1995. Night cultures, night economies. *Planning Practice and Research*, 10 (2), 121–126. doi:10.1080/02697459550036667

Bromley, R.D.F., Tallon, A.R., and Thomas, C.J., 2003. Disaggregating the space - time layers of city-centre activities and their users. *Environment and Planning A*, 35 (10), 1831–1851. doi:10.1068/a35294

Calabrese, F., *et al*., 2011. Real-time urban monitoring using cell phones: a case study in Rome. *IEEE Transactions on Intelligent Transportation Systems*, 12 (1), 141–151. doi:10.1109/TITS.2010.2074196

Cipriani, R., 2013. The many faces of social time: a sociological approach. *Time & Society*, 22 (1), 5–30. doi:10.1177/0961463X12473948

Clemence, G.M., 1959. Ephemeris time. *Astronomical Journal*, 64, 113–115.

Csáji, B.C., *et al*., 2013. Exploring the mobility of mobile phone users. *Physica A: Statistical Mechanics and its Applications*, 392 (6), 1459–1473. doi:10.1016/j.physa.2012.11.040

Debnath, A.K., *et al*., 2014. A methodological framework for benchmarking smart transport cities. *Cities*, 37, 47–56. doi:10.1016/j.cities.2013.11.004

Feldman, L.P. and Hornik, J., 1981. The use of time: an integrated conceptual model. *Journal of Consumer Research*, 7 (4), 407–419.

Filion, P., 2000. Balancing concentration and dispersion? Public policy and urban structure in Toronto. *Environment and Planning C: Government and Policy*, 18 (2), 163–189. doi:10.1068/c2m

Forsyth, A., *et al*., 2007. Does residential density increase walking and other physical activity? *Urban Studies*, 44, 679–697. doi:10.1080/00420980601184729

Griffiths, P., *et al*., 2010. Mapping megacity growth with multi-sensor data. *Remote Sensing of Environment*, 114 (2), 426–439. doi:10.1016/j.rse.2009.09.012

Hägerstrand, T., 1970. What about people in regional science? *Papers of the Regional Science Association*, 24 (1), 6–21. doi:10.1007/BF01936872

Hancke, G.P., De Carvalho E Silva, B., and Hancke Jr, G.P., 2013. The role of advanced sensing in smart cities. *Sensors*, 13, 393–425. doi:10.3390/s130100393

Handy, S.L., Cao, X., and Mokhtarian, P.L., 2005. Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, 10 (6), 427–444. doi:10.1016/j.trd.2005.05.002

Jaagus, J. and Ahas, R., 2000. Space-time variations of climatic seasons and their correlation with the phenological development of nature in Estonia. *Climate Research*, 15 (3), 207–219. doi:10.3354/cr015207

Järv, O., Ahas, R., and Witlox, F., 2014. Understanding monthly variability in human activity spaces: a twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*, 38 (1), 122–135. doi:10.1016/j.trc.2013.11.003

Kwan, M.-P. and Neutens, T., 2014. Space-time research in GIScience. *International Journal of Geographical Information Science*, 28 (5), 851–854. doi:10.1080/13658816.2014.889300

Lambiotte, R., *et al*., 2008. Geographical dispersal of mobile communication networks. *Physica A: Statistical Mechanics and its Applications*, 387, 5317–5325. doi:10.1016/j.physa.2008.05.014

Lee, J. and Lee, H., 2014. Developing and validating a citizen-centric typology for smart city services. *Government Information Quarterly*, 31, S93–S105. doi:10.1016/j.giq.2014.01.010

Lovatt, A. and O'Connor, J., 1995. Cities and the night-time economy. *Planning Practice and Research*, 10 (2), 127–134. doi:10.1080/02697459550036676

Meijers, E., 2007. From central place to network model: theory and evidence of a paradigm change. *Tijdschrift Voor Economische En Sociale Geografie*, 98 (2), 245–259. doi:10.1111/tesg.2007.98.issue-2

Panda, S., Hogenesch, J.B., and Kay, S.A., 2002. Circadian rhythms from flies to humans. *Nature*, 417, 329–335. doi:10.1038/417329a

Positium, L.B.S. 2014. Feasibility study on the use of mobile positioning data for tourism statistics, Consolidated Report Eurostat Contract No 30501.2012.001-2012.452, 31 p. Eurostat. Available from: http://ec.europa.eu/eurostat/web/tourism/methodology/projects-and-studies [Accessed 20 May 2015].

Pred, A., 1981. Social reproduction and the time-geography of everyday life. *Geografiska Annaler Series B-Human Geography*, 63, 5–22.

Ripley, B.D., 2004. *Spatial statistics*. Hoboken, NJ: Wiley-Interscience.

Roenneberg, T., Wirz-Justice, A., and Merrow, M., 2003. Life between clocks: daily temporal patterns of human chronotypes. *Journal of Biological Rhythms*, 18, 80–90. doi:10.1177/0748730402239679

Silm, S. and Ahas, R., 2010. The seasonal variability of population in Estonian municipalities. *Environment and Planning A*, 42 (10), 2527–2546. doi:10.1068/a43139

Silm, S. and Ahas, R., 2014. The temporal variation of ethnic segregation in a city: evidence from a mobile phone use dataset. *Social Science Research*, 47, 30–43. doi:10.1016/j.ssresearch.2014.03.011

Sobolevsky, S., *et al*., 2013. Delineating geographical regions with networks of human interactions in an extensive set of countries. *PLoS ONE*, 8 (12), e81707.

Sorokin, P.A. and Merton, R.K., 1937. Social time: a methodological and functional analysis. *American Journal of Sociology*, 42 (5), 615–629.

Yuan, Y. and Raubal, M., 2012. Extracting dynamic urban mobility patterns from mobile phone data. *In*: N. Xiao, M.-P. Kwan, M. Goodchild, and S. Shekhar, eds., *Proceedings of seventh international conference on geographic information science (GIScience)*, 18–21 September, Columbus, OH. Lecture Notes in Computer Science. Berlin: Springer, 354–367.

Yuan, Y. and Raubal, M., 2014. Measuring similarity of mobile phone user trajectories– a spatio-temporal edit distance method. *International Journal of Geographical Information Science*, 28 (3), 496–520. doi:10.1080/13658816.2013.854369

Yuan, Y., Raubal, M., and Liu, Y., 2012. correlating mobile phone usage and travel behavior - a case study of Harbin, China. *Computers, Environment and Urban Systems*, 36 (2), 118–130. doi:10.1016/j.compenvurbsys.2011.07.003