Using city-bike stopovers to reveal spatial patterns of urban attractiveness

Krystian Banet\textsuperscript{a} Vitalii Naumov\textsuperscript{b} and Rafał Kucharski\textsuperscript{c}
\textsuperscript{a}Department of Transportation Systems, Cracow University of Technology, Krakow, Poland, ORCiD: 0000-0002-9560-483X;
\textsuperscript{b}Department of Transportation Systems, Cracow University of Technology, Krakow, Poland, ORCiD: 0000-0001-9981-4108;
\textsuperscript{c}Department of Transport & Planning, Delft University of Technology, Delft, the Netherlands, ORCiD: 0000-0002-9767-8883

ABSTRACT
We demonstrate how digital traces of city-bike trips may become useful to identify urban space attractiveness. We exploit their unique feature – stopovers: short, non traffic-related stops made by cyclists during their trips. As we demonstrate on the case-study of Kraków (Poland), when applied to a big dataset, meaningful patterns appear, with hotspots (places with long and frequent stopovers) identified at both the top tourist and leisure attractions as well as emerging new places.

We propose a generic method, applicable to any spatiotemporal city-bike traces, providing results meaningful to understand both the general urban space attractiveness and its dynamics. With the proposed filtering (to mitigate a selection bias) and empirical cross-validation (to rule-out false-positive classifications) results effectively reveal spatial patterns of urban attractiveness. Valuable for decision-makers and analysts to enhance understanding of urban space consumption patterns by tourists and residents.

KEYWORDS
spatial-data; city-bike; bike-sharing system; tourist hotspots; digital footprints; urban space; tourist attractiveness

1. Introduction

Identifying urban places attractiveness and quantifying it is of high importance for policymakers, who can better design a city for city users; for the users, who may know which places are attractive; and for the local economy, which can find the optimal locations for their businesses. Yet urban space attractiveness is not at all easy to define, delimit and quantify. Cities are used by various groups, from daily commuters, through local visitors, business travellers, to tourists. Each with various activity patterns, needs and perceptions of urban space attractiveness. Collectively creating a complex spatial patterns, dynamically changing with emerging trends and fashions.
In this study we demonstrate how a big spatiotemporal datasets of mobility traces may be used as a proxy revealing attractiveness of urban spaces. We contribute to the research stream where large sets of aggregated digital footprints are analysed to provide novel insights into how people experience the city (Girardin, Vaccari, Gerber, Biderman, & Ratti, 2009). By utilizing a relatively less exploited dataset (city bike traces) and exploring its unique feature (so-called stopovers) we reveal meaningful and valuable spatiotemporal patterns.

1.1. Literature review

In this section we first introduce the notions of urban attractiveness for tourists and local users along with methods to quantify and measure it. Then, we review a variety of recent methods leveraging on big datasets of digital footprints and their application to urban attractiveness. Finally we discuss city-bike systems, and a unique feature of digital-footprints left by city-bike users - stopovers.

1.1.1. Urban space attractiveness

Following the definition of (Biernacka & Kronenberg, 2018), the urban space is attractive, when one willingly wants to use it and spend her/his time there, and when this space corresponds with one’s individual needs, expectations and preferences. Attractiveness of urban space is not at all easy to define, delimit and quantify (Boivin & Tanguay, 2019), with substantially different perception of urban space for tourists and locals (Kianicka, Buchecker, Hunziker, & Müller-Böker, 2006), notwithstanding both user groups are now better understood thanks to recent studies. For instance, through the indicators to measure urban quality proposed by Garau and Pavan (2018), or with ‘City Love Index’, lately introduced by Kourtit, Nijkamp, and Wahlström (2020), which pinpoints attractiveness characteristics based on perceptions of urban quality by residents and their affinity with city life. Residents’ urban space consumption is associated mainly with their daily activities (Gonzalez, Hidalgo, & Barabasi, 2008), however its attractiveness for leisure purposes becomes increasingly significant (Thees, Zacher, & Eckert, 2020), better understood (Johnson & Glover, 2013) and quantified (Biernacka, Kronenberg, & Laszkiewicz, 2020).

Likewise, the tourists’ behaviour is better understood (for a thorough review we refer to Cohen, Prayag, and Moital (2014)) through studies where various segments (Stangl, Prayag, & Polster, 2020), activity-based profiles (Fieger, Prayag, & Bruwer, 2019), socio-demographic groups (Md Khairi, Ismail, & Syed Jaafar, 2019) or groups with specific needs (Lee & King, 2019) are identified. By means of tourist surveying (Jacobsen, Iversen, & Hem, 2019), stated preference experiments (González, Román, & Ortúzar, 2019), or semi-structured interviews (Kianicka et al., 2006) attractiveness is typically related to a set of site-specific attributes (Estiri, Heidary Dahhoe, Hosseini, & Khajehian, 2020) or individual visitors’ perceptions (Cracolici & Nijkamp, 2009; Lee & King, 2019). Which, in turn, allows for a refined notion of tourism attractiveness at a national (Mitra, 2020), regional (Cracolici & Nijkamp, 2009), city (Van der Ark & Richards, 2006), or site (Jacobsen et al., 2019) level. Which however, becomes challenging when within-urban attractiveness needs to be delimited (Zhu, 2020).

While attractiveness at macro-level can be identified via surveys, observing tourist movements plays a fundamental role in understanding their behaviour within the urban space (McKercher & Lau, 2008). To this end, tourists’ spatiotemporal behaviour – encompassing trajectories (movements between activities) (Zakrisson & Zillinger,
and stops (either at attractions, or to eat, rest, do shopping, etc.) (Caldeira & Kastenholz, 2020) - is analysed with implicit assumption that, in general, consumers of urban space spend more time in attractive spaces (Gehl, 2011).

In such context, the movement along the multi-attraction itinerary can be observed (Huang et al., 2020), with participation time (Caldeira & Kastenholz, 2020) or time spent per bloc (Espelt & Benito, 2006) used as attractiveness intensity indicators. Early attempts to track tourists’ movements using mental maps or self-completion diaries and surveys were usually time consuming and thus applied only on small sample sizes (Keul & Kühberger, 1997; Thimm & Seepold, 2016).

1.1.2. Digital footprints

Digital footprints are now available in big volumes from numerous sources [J. Li, Xu, Tang, Wang, & Li, 2018] which, coupled with a new kind of tourist that is avid for online content and predisposed to share information on social media, allows for a better understanding of tourist behaviour regarding their spatial distribution in urban destinations [Encalada, Boavida-Portugal, Ferreira, & Rocha, 2017].

Big volumes of data and its high availability seem to overweight limitations, mainly inherent selection bias [Salas-Olmedo, Moya-Gómez, García-Palomares, & Gutiérrez, 2018]. Consequently, big data in smart tourism [Y. Li, Hu, Huang, & Duan, 2017] contributes to understanding spatial patterns around urban tourist destinations and, for instance, to differentiate the overcrowded places from those with the potential to grow, allowing decision-makers to revisit planning and managing towards a sustainable ‘smart’ future.

User-generated social media content (photos on Twitter, Instagram, Flickr, etc. or recommendations and reviews on TripAdvisor or Booking) have been widely explored in numerous studies [Giglio, Bertacchini, Bilotta, & Pantano, 2019; Hasnat & Hasan, 2018; D. Li, Zhou, & Wang, 2018; Miah, Vu, Gammack, & McGrath, 2017; Önder, Koerbitz, & Hubmann-Haidvogel, 2016]. Recently Martí, García-Mayor, and Serrano-Estrada (2021) used Instagram data to reveal detailed picture of urban areas with most tourism-related activity – i.e. sightseeing, shopping, eating and nightlife - in Spanish cities.

While such data may reveal spatial patterns, it does not track the tourist movements, which requires geo-location data from personal devices (mobile phones) or vehicles (e.g. rental cars, scooters or bikes). GPS traces were used e.g. by Girardin et al. (2009) to provide insights into the attractiveness of urban space in NYC, by Orellana, Bregt, Ligtenberg, and Wachowicz (2012) to explore visitor movement patterns in the Dwingelderveld National Park, by Smallwood, Beckley, and Moore (2012) to understand distance decay in destination choice. Zheng, Huang, and Li (2017) used GPS to predict next destination within a Summer Palace in Beijing and Ferrante, De Cantis, and Shoval (2018) tracked cruise passengers at the destinations.

1.1.3. City-bike mobility traces

Lately, bicycle sharing has become an increasingly popular around the world, making usage datasets big enough to study urban dynamics and aggregate human behaviour (Froehlich, Neumann, & Oliver, 2009). City bike systems store records of trips with their origins, destination, and travel times in publicly available big databases, which allows for a rich understanding of mobility patterns (Cantelmo, Kucharski, & Antoniou, 2019). Number of recent studies have used city-bike data e.g. to identify potential
locations for new stations, estimate bicycle flows and usage, understand social and demographic context or predict usage in real-time (Caulfield, O’Mahony, Brazil, & Weldon, 2017; Eren & Uz, 2020; Frade & Ribeiro, 2014; Imani, Eluru, El-Geneidy, Rabbat, & Haq, 2014; Salon, Conway, Wang, & Roth, 2019; Tran, Ovtracht, & D’Arcier, 2015; K. Wang & Akar, 2019). Parallel studies investigate how bike trips are affected by urban space factors such as: the number of retail stores and business offices near bike stations (Lin et al., 2020), demographic features (J. Wang & Lindsey, 2019) or land-use (Kutela & Teng, 2019).

How city-bikes are used by tourists was also studied. (Vogel, Greiser, & Mattfeld, 2011) identified that stations dominating between noon and afternoon were located directly near tourist hotspots, (Brinkmann, 2020) has shown differences in city bike usage between tourists and frequent users in Rio de Janeiro and Miami Beach. (Buning and Lulla, 2020) revealed different usage patterns between local residents and visitors, showing that visitors primarily use city-bike for leisure-based urban exploration, while residents' use bikes mainly to commute. However, up to our knowledge, the unique feature of city-bike traces - stopovers was not exploited so far.

1.2. Study overview

Core of the proposed method lays in the concept of a stopover, a short stop made by a city-bike user during her/his trip. Bicycle is not returned at the station, but stays with the user during a stopover. Stopovers are typically short, since for longer stops users typically return bicycle to the docking station due to time-based fare scheme. Stopover is not related to traffic, as we explicitly filter traffic-related stops e.g. at traffic lights. With such notion of stopover we may limit it to non-commuting trips, since commuters rent bikes to quickly reach the destination rather than to have stopovers.

Users may stopover for variety of reasons. The actual interpretation of stopover depends on the user type. We introduce city-bike users typology on figure 1. Since the user details are missing we cannot distinguish commuters (using city-bike to reach their workplace), from local recreational users (having a weekend tour around green areas of the city), from business visitors (using city bike to reach the dinner with a client) and tourists (using city-bike to visit recommended tourist destinations).

Nonetheless, we hypothesise that if city-bike user stops where she/he does not need to, it is mainly due to the place attractiveness, which may be either touristic, recreational, commercial or of any other kind. We further hypothesise that places where stopovers are frequent and long (following Gehl (2011)), denoted urban hotspots, are attractive.

While we argue that identified hotspots are meaningful and valuable, we refrain from naively interpreting them as attractive urban hotspots. Like any other automated classification method, accuracy of our method is not perfect, as we illustrate with confusion matrix on fig.2a. The wrong classifications are either when our method failed to identify actually attractive places (e.g. places not accessible with bike, or outside of city-bike system) or when it identified places which are not attractive (where stopovers were not due to attractiveness, but for other reasons). We argue that validation of our results is straightforward, since each identified place may be examined against its true attractiveness relying either on the expert knowledge, other sources (social media or digital footprints from other sources) or a field visit.
Figure 1. City bike users classification. While local commuters stop mainly due to traffic, other user groups are likely to stopover during their trip. Both visitors and local non-commuters may stop at attractive leisure and touristic places, as well as to supply some their needs (e.g. shopping). Places with frequent and long stopovers (hotspots), after empirical validation, may be used as a proxy of urban space attractiveness.

Figure 2. Accuracy of our classification against actual attractiveness (left) and classic static methods (right).
Figure 3. Urban space attractiveness in Kraków, Poland. Identified hotspots were not only the classic points of any tourist itinerary, but also emerging places not mentioned in travel guides, e.g.: 1. the Dolnych Młynów pubs and clubs 2. Karmelicka str. gardens, 3. the Dąbski Pond.

1.3. Research problem, gap and contribution

We contribute to the stream of research aiming to reveal spatial patterns of urban attractiveness. Identifying tourist hotspots and determining their dynamically changing attractiveness level is of crucial importance for decision-makers, who can now design the city better with attractiveness perception of residents and tourists in mind. Thus, in rapidly changing urban landscape, we need a dynamic method adapting to the recent trends and fashions of tourists (Dunne, Flanagan, & Buckley, 2011) and residents (Kourtit et al., 2020). Classic methods relying on expert knowledge and costly surveying fail to provide the detailed picture and are inherently static. Recently, number of methods were proposed where big volumes of digital footprints were applied to reveal behaviour of residents and tourists in urban areas. Consequently, the delimitation of attractive urban spaces has become more detailed and underpinned with users’ behaviour (observed via their digital footprints) and perception (understood thanks to surveying). The objective of this paper is to demonstrate how this picture can be improved by using a new source of data and its unique features.

In this paper we exploit the potential of stopovers to reveal the spatial patterns. We hypothesise that the stopovers are related to the space attractiveness and verify it on a case-study. Nonetheless, observing stopovers is challenging. Stopovers cannot be read from social media data, even geotagged, which does not provide a participation time and, since it requires users action, content is not posted from all the places perceived as attractive. Detailed spatiotemporal digital traces are needed to reveal stopovers, and only the active travel modes (walking, scooters, bicycles, etc.) allow for unrestricted stopovers. Cars are used by urban space consumers to a limited extend and most of attractive places are not accessible with car. Cars can be traced only to their parking spot and public transport passengers up to their bus stop.

Pedestrians exploring urban space are the least restricted to make spontaneous
stopovers. However tracing pedestrians typically raises privacy issues and big volumes of personal mobile location data are not easily available. The privacy issues are partially overcome in a station-based systems, which does not contain sensitive personal data. Consequently, the stopovers may be easily observed on a large scale only from the shared systems like scooters and city-bikes, which is their unique feature. In this study we exploit its potential.

Mining stopovers from detailed trajectories is not trivial. To this end we propose a novel method which allows first to identify stopovers from spatiotemporal trajectories and then to filter stopovers clearly not related to attractiveness. While results of the method needs to validated against external sources and local knowledge (as illustrated on fig.2) the revealed pattern accurately reproduced tourist hotspots of Kraków. While proposed method works with unlabelled data, the results may be refined when extra labels (sociodemographics or user-type) are available.

Applying our method to the case of 35 thousands traces from Kraków, revealed a surprisingly meaningful and correct spatial pattern (fig.3). Not only the main tourist attractions were properly identified, but also other insightful findings appeared. We identified a number of hidden gems, known only to locals, as well as newly emerging places, recently gaining popularity and often not yet listed on tourist websites. Such places are unlikely to be timely identified via static studies, relying on expert knowledge (like in Faracík et al. (2008)), or surveys (like Kianicka et al. (2006)) as we illustrate on fig.2), which is a central contribution of the paper. We demonstrate this with three examples on fig.3.

The study contributes to solving the problem of the tourist hot-spots identification by using the exact statistical data against the traditional approaches that infer based on unreliable information from the surveys and opinions of social media users. Identifies a new feature of well-known digital footprint source, stopovers, widely available only in city-bike traces.

The paper is organised as follows. In the next section we introduce a generic method to identify stopovers in mobility traces and apply it to the city bike datasets. We introduce a set of filters to calibrate the method before we synthesise the data on the spatial grid. In section 3 we illustrate the method using the example of Kraków, where stopovers identified in 35,000 bike trips yielded a grid that was validated against actual tourist hotspots. Finally, in section 4, we synthesise the results and discuss the potential applications and limitations of the proposed method.

2. Method

We first formalize how stopovers are identified in the raw dataset, followed with a filtering rules, after which only meaningful stopovers remain in the dataset. Consequently, we aggregate the stopovers over a spatial grid and classify cells into four levels of attractiveness, the outcome of the method. The code to read the data from gpx-files, identify stopovers and compute attractiveness grid is publicly available.

2.1. Stopovers

We analyse trip tracks, i.e., chronologically ordered sets of track-points:

\[ \text{https://github.com/naumovvs/city-bikes-analysis} \]
\( Track = \{TP_i\}, \quad i = 1, \ldots, N_{TP} \) \hspace{1cm} (1)

, where each track point \( TP \) is defined as the time \( t \) and position:

\[ TP = \langle t, lon, lat \rangle \] \hspace{1cm} (2)

For convenience, we use a dual definition, where trip becomes a set of trip segments:

\[ Track = \{TS_j\}, \quad j = 1, \ldots, N_{TS} \] \hspace{1cm} (3)

and \( TS_j \) is the \( j \)-th trip segment of the journey track defined by the couple of neighbouring track points \( TP_o \) and \( TP_d \):

\[ TS = \{TP_o, TP_d\} \] \hspace{1cm} (4)

From each trip segment we read: \( t_s \) - the travel duration for the trip segment [h]; \( d_s \) - the distance, calculated with haversine formula [km] and \( v_s \) - the average travel speed, defined as the distance \( d_s \) divided by the travel duration \( t_s \) [km/h]. Consequently, the raw trips data now becomes:

\[ Trip = \langle ID, Track, t_{tr}, t_{idle}, d \rangle \], \hspace{1cm} (5)

where \( ID \) is the unique number identifying a trip in the dataset; Track is the reference to the object representing the GPS track as a set of track segments; \( t_{tr} \) is the total travel time according to the track points data (the difference between the time moments when the last and the first track points in the track were read) [h]; \( t_{idle} \) is the total idle time during the journey [h]; \( d \) is the travel distance [km].

The total idle time is defined as the sum of travel durations for those travel segments for which the location has not been changed:

\[ t_{idle} = \sum_{TS_{idle}} t_s, TS_{idle} = \{TS_j : d_j = 0\}, \quad j = 1, \ldots, N_{TS} \] \hspace{1cm} (6)

\( TS_{idle} \) is a set of all segments within the trip that have zero travel distance. Stopovers, central element of the proposed method, are a set of consequent travel segments with null distance:

\[ Stopover = TS_{idle}\{TS_k, TS_{k+1}t_s\}, \quad 1 \leq k \leq N_{TS}. \] \hspace{1cm} (7)

Note that, according to the definition, more than one stopover could be defined within a trip, if travel segments with zero distance are not consequent segments. We illustrate three selected rides with various number of stopovers on fig.4.
2.2. Filtering

To make sure that identified stopovers are meaningful, we applied the following pre-processing sequence:

(1) The dataset was first cleansed of corrupt records related to signal failures in GPS transmitters. Trip data were eliminated if the GPS outage lasted at least five minutes.

(2) After the first filtering stage, the sample continued to contain many trips with an average speed of 0 km/h. Therefore, we decided to eliminate all trips with a null duration or distance, which were most likely related to situations where a bike was unlocked, but not taken out of the stand, and then locked again, e.g. because of a technical problem.

(3) Subsequently, we removed very short trips, where a technical problem was discovered soon after the bike was rented and the user decided not to continue with the trip. Based on (Naumov & Banet, 2020), a 50 m threshold was used.

On such filtered sample we identified stopovers using the method proposed above. For each recorded trip, stopovers were identified with (eq. 7) with their location and duration. Yet, a heatmap visualization of obtained stopovers revealed a need for further filtering. To this end we applied a second stage of filtering using spatial metadata from OpenStreetMap as follows:

(1) Bike rental/return stops.
Most evident was the need to filter stopovers in the proximity of BSS stations, where what our method identified as stopovers were in fact unlocking and locking the bike and checking its technical condition. We identified threshold of 7m around BSS station to efficiently eliminate trip starts and ends from stopovers. We explored the cumulative number of stops as a function of distance from origins and found a natural cut-off point at 7 meters, after which the number of stops stabilized (Naumov & Banet, 2020).

(2) Stops at traffic lights.
Obviously, most of times when city bike users stop is not for sight-seeing, but at traffic. This had to be filtered with care to obtain meaningful results. Importantly, in the vicinity of most of Kraków’s tourist attractions there are no traffic lights,
so we could safely assume that stopovers around traffic lights are not due to attractiveness. Our analysis revealed that the number of stops stabilized at a radius of 30 m; adopting a greater radius would lead to discarding stopovers unrelated to the presence of a pedestrian crossing or intersection.

(3) Railway crossings stops.
Despite there is just handful of single-level railway crossings in Kraków, stopovers in their vicinity (definitely non-attractive places) biased the emerging picture. Since those were just few points, it was easy to identify them and manually filter at 30m threshold.

(4) Short, traffic related stopovers.
While above filters were spatial, we decided to apply also a temporal filter, which we found efficient in filtering short, traffic related stops. Namely, we found that vast majority of stopovers below 30s were around unsignalized pedestrian crossings. So we filtered stopovers in vicinity of pedestrian crossings shorter than 30s.

After above stages of filtering, the meaningful spatial patterns started to emerge, with heatmaps now clearly resembling tourist attractiveness, rather than a traffic map. Notably, in the above we did not need to map-match traces, which makes the methods light and generally applicable.

2.3. Aggregation
For meaningful and quantifiable visualization, we divided the analysed area into a number of rectangular fields of a given size $S$. We used a spatial grid that can be represented as the following matrix:

$$\text{Grid} = \{Field_{ij}\}^{S \times S}, \quad i = 1 \ldots S, \quad j = 1 \ldots S,$$

where $Field_{ij}$ is the rectangular field representing the part of the analysed area; $S$ is the grid size (the greater number of cells, the more detailed results). For each field we get:

$$\text{Field} = \{n_{st}, t_{st}, t_F\},$$

where $n_{st}$ is the total number of stopovers in the field; $t_{st}$ is the total duration of all the stopovers in the field $\{h\}$, and $t_F$ is the mean stopover time (calculated only for cells with more than three records):

$$t_F = \begin{cases} 
0 & n_{st} < 3 \\
\frac{t_{st}}{n_{st}} & \text{otherwise} 
\end{cases}$$

All being potentially useful to reveal the spatial attractiveness. As we demonstrated for our case-study in the next section, we decided to base attractiveness on the mean duration of stopovers rather than their number or total duration. To sharpen the emerging picture, we classified grid cells into four attractiveness classes, representing quartiles of mean stopover duration ($t_F$), where 4-th class represents highest attractiveness and 0-th lowest. The final outcome of the method is a spatial grid with rank (from 0 to 3) of urban space attractiveness for each field.
Table 1. Number of stopovers and their statistics in the subsequent filtering stages.

| stage | #stops | mean [s] | std [s] | min [s] | 25% [s] | 50% [s] | 75% [s] | max [s] | sum [h] |
|-------|--------|----------|--------|--------|--------|--------|--------|--------|--------|
| raw   | 54143  | 79.17    | 291.63 | 1      | 15     | 25     | 65     | 27350  | 1190.77|
| 1st   | 9639   | 280.93   | 559.30 | 1      | 85     | 90     | 255    | 8290   | 752.20 |
| 2nd   | 6277   | 335.83   | 641.13 | 1      | 85     | 155    | 275    | 8290   | 585.56 |
| 3rd   | 6219   | 337.74   | 649.78 | 1      | 85     | 160    | 275    | 8290   | 583.45 |
| 4th   | 5791   | 361.80   | 667.09 | 31     | 85     | 170    | 335    | 8290   | 582.00 |

3. Results

We illustrate the proposed method with the case of Kraków, Poland, one of emerging tourism centres in Europe. With its rich history and unique cultural heritage, the city has attracted a growing number of tourists in the last decade. For Kraków, tourism is not just an important source of revenue, but also a major social phenomenon that shapes its urban identity. Tourism in Kraków has long been concentrated around the historic city centre (Old Town, Wawel Castle, Kazimierz), i.e. the urban complex inscribed in the UNESCO World Heritage List. Yet now it extends to neighbouring areas, such as Kleparz, Krowodrza, Zabłocie, Stare Podgórze, and Nowa Huta, due activities aimed at decentralizing tourist traffic [Tracz & Semczuk, 2018]. Visitors are becoming more and more heterogenic, spanning from John Paul II related pilgrims, to city-breakers focused on nightlife, from high-school pupils visiting their national royal Castle for the first time, to frequent visitors looking beyond the top-10 sights. These dynamics and diversity yield rapid and complex patterns which are hard to trace and quantify.

We used the data from the local city bike system, ‘Wavelo’, established in 2008 and gradually expanded until 2019 when its area covered majority of the city. The records covered one week of the high tourist season of 2017, i.e., from 31st of May to 7th of June, when weather was auspicious for bike traffic and recreation, with mean temperatures between 16 and 21°C and barely any rainfalls. The dataset composed of 34,969 tracks of Wavelo city bike users was sufficient to obtain clear and meaningful patterns. The high-fidelity and pre-processed mobility traces in the gpx-format were obtained from GPS transmitters attached to every Wavelo bike.

The first step was to cleanse the dataset provided by the city bike system. First, we eliminated all data corrupted by GPS transmission failures (a total of 5,946 trips). At the second stage, 40 trips with a null duration and 635 trips with a null distance were removed. The third stage, which involved filtering out short trips, identified 421 trips with a distance shorter than 50 m. Once these were eliminated, the final sample consisted of 27,927 routes.

The number of stopovers in the cleansed sample was 54,143, with a mean stopover time of ca 80s. After the last filtering stage, the number of stopovers dropped to 5,791, while their mean duration rose to a little over 6 minutes; only 25%, however, were longer than 5 minutes 35 seconds (tab. 1). The largest drop in the number of stopovers was recorded after the first step, which involved eliminating those within a radius of 7 m from the trip’s origin. Further decreases were less steep, but a clear relationship could still be observed between the increase of the mean duration and the lower number of stopovers in the sample.

After each filtering stage, we visualised the dataset for verification of the obtained
results. The unfiltered sample was dominated by punctual stops near rental stations. The first filtering stage allowed us to identify stops that were not related to the trip origin or destination. After the first stage, the map still contained punctual hotspots related to stopovers at junctions and level crossings which were successfully filtered with subsequent filtering stages. Finally, most stopovers were identified in the touristic and leisure places like Vistula Boulevards, Old Town and the Main Market Square.

In order to quantify results, we created a grid where each field was assigned the corresponding number of stopovers and total stopover time (eq. 9). In the case at hand, the area delimited by the outer geographical coordinates of the recorded stopovers was divided into 10,000 fields of the same geographical longitude and latitude, i.e., $0.0015^\circ \times 0.0036^\circ$. Figure 5 shows values on the grid of: number of stops (a), total stopover time (b), mean stopover time (c) and mean stopover time in four classes (d). Based on the emerging patterns, it was evident that the mean stopover time provides a meaningful proxy to identify urban hotspots. Other ones yielded both false-positive as well as true-negative errors, where identified hotspots were not attractive and attractive hotspots were not identified, respectively.

Most fields in the attractiveness identification grid, i.e., 94.74%, have a rating of 0, but the most attractive areas of the city, such as the Old Town, the Vistula Boulevards, the Benedictine Abbey in Tyniec, the Kolna kayaking centre, the Polish Aviators’ Park, Bagry Lake or Nowa Huta Lake, achieved a high attractiveness score. Table 2 shows
the number of fields in each attractiveness category and their percentage share in the total number of fields in the attractiveness identification grid.

### 3.1. Validation

Since the proposed method is explorative and aims for more complete identification of previously unrevealed attractive urban hotspots, its validation is not straightforward. Nonetheless, to demonstrate its capabilities we validated our results against typical sources of tourist attractiveness. Faracik et al. (2008) evaluated urban space in terms of its tourism attractiveness, which was later adopted as an official and the latest tourism policy by the Mayor of Kraków. In the comprehensive study, they relied on their expert knowledge to select the natural, cultural, accommodation and services factors in the assessment process and, similarly to our study, classified city space into four classes, as shown in fig. 7.

Most of the hotspots identified by our method overlap with those mentioned in the literature. Figures 8 to 11 zoom in the attractiveness grid in selected areas of Kraków and discuss them. By comparing figure 7 with figures 8 to 11 one can see greater details of hotspots locations and complex, yet clear patterns resulting from our method.

In table 3 we compare most important official tourist attractiveness (highly ranked in Faracik et al. (2008)) with our method results. Our method managed to correctly identify all the tourist attractions from the official sources of attractiveness (compare our findings on fig 6 with official attractiveness on fig 7), with two exceptions, both poorly accessible by bike.

First one is pilgrims centre in Łagiewniki (south), which is typically visited by elder tourists who rarely use city bikes and is poorly accessed by bicycle (as it is located on the hill). The second was the Wolski Forest, with the Zoo and the Pilsudski Mound. While considered as one of the most attractive spots in the city, for topographical reasons, this place is popular among mountain bikers rather than Wavelo users.

On the other hand, some places classified as attractive with our method, were clearly not touristic shopping malls. Shopping can be perceived attractive by tourists and locals, several of shopping malls in Kraków are the attractive ones. They are located near the Old Town (Galeria Krakowska) and Kazimierz (Galeria Kazimierz). However the two examples that we use: the Krokus mall, and shopping centre in Norymberska Street are clearly not attractive and used by locals to supply their basic needs. Our method failed to filter them out, yet manual post-processing with a basic background field-knowledge allowed to effectively interpret such false-positive cases.

| class | # fields | share  |
|-------|----------|--------|
| 0 (lowest) | 9474 | 94.74% |
| 1     | 183     | 1.83%  |
| 2     | 183     | 1.83%  |
| 3 (highest) | 160 | 1.60%  |

Table 2. Attractiveness classes of attractiveness identification grid fields.
Figure 7. Four classes of tourist attractiveness in the official policy of Mayor of Kraków [Faracik et al. 2008].

| Place                                                  | rank:          |           | our accuracy* | comment                        |
|--------------------------------------------------------|----------------|------------|---------------|--------------------------------|
|                                                        | official       | our        |               |                                |
| **Central (fig. 8)**                                   |                |            |               |                                |
| the Main Market Square                                 | 3              | 3          | TP            |                                |
| Wawel Castle, and the Vistula Boulevards               | 3              | 3          | TP            |                                |
| Kazimierz quarter                                     | 3              | 3          | TP            |                                |
| the city beach in the Courland Boulevard               | -              | 3          | TP            | recently opened                |
| Blonia and the Rudawa Valley                           | 2              | 3          | TP            |                                |
| the Jordan’s Park                                      | 2              | 3          | TP            |                                |
| **South-west (fig. 9, 10)**                            |                |            |               |                                |
| Cricoteka                                              | 1              | 3          | TP            |                                |
| Schindler’s Factory and the MOCAK                      | 2              | 3          | TP            |                                |
| the Krakus Mound                                       | 2              | 3          | TP            |                                |
| Zoo and Pilsudski Mound                                | 2              | 0          | TN            | poorly accessible by bike      |
| Bagry Lake                                             | 3              | 3          | TP+           | primarily for locals           |
| the Vistula Boulevards in Stare Dębniki,               | 2              | 3          | TP            |                                |
| Pilgrim center Lagiewniki                              | 3              | 0          | TN            | poorly accessible by bike      |
| the Vistula Boulevards in Ludwinów                     | 2              | 3          | TP            |                                |
| Zakrzówek                                              | 1              | 3          | TP+           | primarily for locals           |
| the kayaking trail in Tyniec                           | 2              | 3          | TP            |                                |
| the Benedictine Abbey in Tyniec                        | 3              | 3          | TP            | distant yet bike-accessible    |
| **East (fig. 11)**                                     |                |            |               |                                |
| Centralny Square                                       | 2              | 3          | TP            |                                |
| the Polish Aviators’ Park                              | 1              | 3          | TP+           |                                |
| Nowa Huta Lake                                         | 2              | 3          | TP            |                                |
| Nowa Huta Meadows                                     | -              | 3          | TP+           | picnic spot for bike trips     |
| Krokus shopping mall                                   | -              | 3          | FP            | hardly not attractive          |

Table 3. Most attractive tourist places according to official listing [Faracik et al. 2008] compared with our classifications. Italics denote places either not identified or wrongly classified as attractive our method. *TP - true positive, TP+ - added-value to the official sources, TN - true negative, FP - false positive, compare with fig. 2.
Figure 8. Selected highest-rated spots in Central Kraków: 1. Błonia with the Rudawa Valley and the Jordan Park, 2. the Old Town with the Main Market Square, 3. Wawel Hill with the Vistula Boulevards, 4. Kazimierz with the Vistula Boulevards, 5. the city beach area, highly popular among locals is the new riverside and not listed in official guides. The old town (2) is now more detailed, making the most attractive spots visible.

Figure 9. Selected highest-rated spots in South Kraków: 1. the kayak rental station, 2. Zakrzówek lake, 3. the Vistula Boulevards in Stare Dębniki, including the Dębnicki Park, 4. the Vistula Boulevards in Ludwinów, 5. the area of Cricoteka and the Podgórski Market Square, 6. the area of Schindler’s Factory and the MOCAK, 7. Bagry Lake. Pilgrims centre Łagiewniki (to the south), highly ranked in official guides, not identified in our method due to low bike accessibility. The shopping mall at Norymeraska street (down from hotspot 2) was identified as attractive, which is clearly a false-positive case that has to be filtered manually.

Figure 10. Selected highest-rated spots in West Kraków: 1. the Benedictine Abbey in Tyniec, 2. the kayaking trail in Tyniec. Both attractive yet distant, which leaves a trace of short breaks along the highly popular bike path stretching by the river between old town and Tyniec.
4. Conclusions and discussion

We proposed a generic method applicable for any spatiotemporal data from city-bikes, which, since city bikes are nowadays present in most of metropolises worldwide, makes it applicable to explore spatial patterns of stopovers in broad range of cities. The light and replicable method, relying on standard spatiotemporal trip tracks allows to identify hotspots - places with frequent and long stopovers. Which, as we demonstrate are a good proxy of the space attractiveness. By assuming that most of stopovers identified with the method are related to tourism and/or leisure, we identified complex and meaningful spatial patterns, clearly pointing towards city’s most attractive urban hotspots. The results shown that the most frequent stopover locations of Wavelo bike users were, in fact, concentrated in the proximity of the most attractive cultural and natural assets. With the proposed automated manual filtering, not relying on local field knowledge, one can reveal number of valuable findings, both in terms of identifying unknown places as well as quantifying well-known ones. While the local knowledge may refine the results and make it reliable indicator of urban attractiveness.

Results of our case-study proved that the method effectively identified most of established Kraków tourist attractions. The recent dynamism in behaviour of tourists, shifting from well-known paths to exploring newly emerging sights, was evident from the emerging patterns. Our method managed to cover it and quantify those changes.

Despite relying on personal and potentially sensitive data, we find the method transparent and ethical. Even though the precise spatial path is recorded, it always starts and ends at the city bike station, rather than at personally sensitive home or workplace. While the users’ ID is not stored in the dataset, his sociodemographic attributes may be disclosed which would enable more detailed analysis differentiating locals from visitors, young from older users, etc. The proposed method has potential in real-time monitoring and can be potentially automated to report the attractiveness and its relative changes over time. Making it an efficient tool for policy makers to monitor shifts in tourist behaviour.

Importantly, in the post-COVID context our method offers an efficient and inexpensive monitoring framework allowing to understand how the pandemic changes impacted the consumption of urban space by tourists, locals and visitors. Allowing to quickly identify most visited, possibly crowded, places where intervention may be needed to stop virus-spreading.
Obviously, our method has some limitations. It relies on sequence of filters, where thresholds need to be manually parameterized (e.g., cut-off distance from traffic lights, or from station). While it effectively filters commuting trips, as long as user type remains unknown, the leisure remains indistinguishable from tourism, and local residents from visitors. This shall be further enhanced with labelled data. While the city bike systems are often limited in space and rental stations are not evenly distributed everywhere in the city. This was not the case for Kraków, yet in cities where coverage is not complete, this may obscure the overall image and fail to yield a complete spatial pattern.

Finally, to verify the results before drawing conclusions a basic field knowledge is needed. In the case of Kraków, some definitely non-attractive shopping malls were misclassified with our method. Yet in any case, virtual or physical site visit may always verify its actual attractiveness. For instance, for us, the locals of Kraków, the data revealed the place that we were not aware of (hidden garden at Karmelicka street).

We believe that with this method we fill some of the gaps in research on the urban space attractiveness for residents and tourists.

**Acknowledgement(s)** We thank the City of Kraków, the operator of Wavelo city bike for the data for analysis.

**References**

Biernacka, M., & Kronenberg, J. (2018). Classification of institutional barriers affecting the availability, accessibility and attractiveness of urban green spaces. *Urban Forestry & Urban Greening, 36*, 22–33.

Biernacka, M., Kronenberg, J., & Łaszkiewicz, E. (2020). An integrated system of monitoring the availability, accessibility and attractiveness of urban parks and green squares. *Applied Geography, 116*, 102152.

Boivin, M., & Tanguay, G. (2019). Analysis of the determinants of urban tourism attractiveness: The case of québec city and bordeaux. *Journal of destination marketing & management, 11*, 67–79.

Brinkmann, J. (2020). Active balancing of bike sharing systems.

Buning, R., & Lulla, V. (2020). Visitor bikeshare usage: tracking visitor spatiotemporal behavior using big data. *Journal of Sustainable Tourism.

Caldeira, A., & Kastenholz, E. (2020). Spatiotemporal tourist behaviour in urban destinations: a framework of analysis. *Tourism Geographies, 22*(1), 22–50.

Cantelmo, G., Kucharski, R., & Antoniou, C. (2019). A low dimensional model for bike sharing demand forecasting. In Mt-its 2019 - 6th international conference on models and technologies for intelligent transportation systems.

Caulfield, B., O’Mahony, M., Brazil, W., & Weldon, P. (2017). Examining usage patterns of a bike-sharing scheme in a medium sized city. *Transportation Research Part A: Policy and Practice, 100*, 152–161.

Cohen, S. A., Prayag, G., & Moital, M. (2014). Consumer behaviour in tourism: Concepts, influences and opportunities. *Current issues in Tourism, 17*(10), 872–909.

Cracolici, M. F., & Nijkamp, P. (2009). The attractiveness and competitiveness of tourist destinations: A study of southern italian regions. *Tourism management, 30*(3), 336–344.

Dunne, G., Flanagan, S., & Buckley, J. (2011). Towards a decision making model for city break travel. *International Journal of Culture, Tourism and Hospitality Research, 3*(2), 158–172.

Encalada, L., Boavida-Portugal, I., Ferreira, C., & Rocha, J. (2017). Identifying tourist places of interest based on digital imprints: Towards a sustainable smart city. *Sustainability
Eren, E., & Uz, V. (2020). A review on bike-sharing: The factors affecting bike-sharing demand. In Sustainable cities and society (Vol. 54, p. 101882). Elsevier Ltd.

Espelt, N., & Benito, J. (2006). Visitors’ behavior in heritage cities: The case of girona. Journal of Travel Research, 44(4), 442–448.

Estiri, M., Heidary Dahooie, J., Hosseini, F., & Khajeheian, D. (2020). Proposing a new model for shopping centre attractiveness assessment by a combination of structural equation modelling (sem. In *and additive ratio assessment (aras). current issues in tourism* (p. 1–19).

Faraci, R., Kurek, W., Mika, M., Pawlusinski, R., Pitrus, E., Ptaszycka-Jackowska, D., . . . Zawiślińska, B. (2008). Waloryzacja przestrzeni miejskiej krakowa dla potrzeb turystyki - raport końcowy.

Ferrante, M., De Cantis, S., & Shoval, N. (2018). A general framework for collecting and analysing the tracking data of cruise passengers at the destination. *Current Issues in Tourism, 21*(12), 1426–1451.

Froehlich, J., Neumann, J., & Oliver, N. (2009). Sensing and predicting the pulse of the city through shared bicycling. In *ijcai international joint conference on artificial intelligence* (Vol. 3, p. 1420–1426).

Garau, C., & Pavan, V. M. (2018). Evaluating urban quality: Indicators and assessment tools for smart sustainable cities. *Sustainability, 10*(3). Retrieved from https://www.mdpi.com/2071-1050/10/3/575

Gehl, J. (2011). *Life between buildings using public spaces*. Island Press.

Giglio, S., Bertacchini, F., Bilotta, E., & Pantano, P. (2019). Using social media to identify tourism attractiveness in six italian cities. *Tourism management, 72*, 306–312.

Girardin, F., Vaccari, A., Gerber, A., Biderman, A., & Ratti, C. (2009). Quantifying urban attractiveness from the distribution and density of digital footprints. *International Journal, 4*, 175–200.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *nature, 453*(7196), 779–782.

González, R., Román, C., & Ortúzar, J. D. (2019). Preferences for sustainable mobility in natural areas: The case of teide national park. *Journal of Transport Geography, 76*, 42–51.

Hasnat, M., & Hasan, S. (2018). Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data. *Transportation Research Part C: Emerging Technologies, 96*, 38–54.

Huang, X., Li, M., Zhang, J., Zhang, L., Zhang, H., & Yan, S. (2020). Tourists’ spatial-temporal behavior patterns in theme parks: A case study of ocean park hong kong. *Journal of Destination Marketing & Management, 15*, 100411.

Imani, A., Eluru, N., El-Geneidy, A., Rabbat, M., & Haq, U. (2014). How does land-use and urban form impact bicycle flows: Evidence from the bicycle-1 sharing system (bixi) in montreal. In *Proceedings of the 93rd annual meeting of the transportation research board*. Washington DC. Retrieved from https://doi.org/10.1002/jgrd.50214

Jacobsen, J. K. S., Iversen, N. M., & Hem, L. E. (2019). Hotspot crowding and over-tourism: Antecedents of destination attractiveness. *Annals of Tourism Research, 76*, 53–66.

Johnson, A. J., & Glover, T. D. (2013). Understanding urban public space in a leisure context. *Leisure Sciences, 35*(2), 190-197. Retrieved from https://doi.org/10.1080/01490400.2013.761922

Keul, A., & Kühlberger, A. (1997). Tracking the salzburg tourist. *Annals of Tourism Research, 24*(4), 1008–1012.

Kianicka, S., bucheccher, M., hunziker, M., & müller-böker, U. (2006). Locals’ and tourists’ sense of place. *Mountain Research and Development, 26*(1), 55–63.

Kourtit, K., Nijkamp, P., & Wahlström, M. (2020). How to make cities the home of people a
‘soul and body’ analysis of urban attractiveness. *Land Use Policy,* 104734.

Kutela, B., & Teng, H. (2019). The influence of campus characteristics, temporal factors, and weather events on campuses-related daily bike-share trips. *Journal of Transport Geography,* 78, 160–169.

Lee, C.-F., & King, B. (2019). Determinants of attractiveness for a seniors-friendly destination: a hierarchical approach. *Current Issues in Tourism,* 22(1), 71–90.

Li, D., Zhou, X., & Wang, M. (2018). Analyzing and visualizing the spatial interactions between tourists and locals: A flickr study in ten us cities. *Cities,* 74, 249–258.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management,* 68, 301–323.

Li, Y., Hu, C., Huang, C., & Duan, L. (2017). The concept of smart tourism in the context of tourism information services. *Tourism Management,* 58, 293–300.

Lin, P., Weng, J., Hu, S., Alivanistos, D., Li, X., & Yin, B. (2020). Revealing spatiotemporal patterns and influencing factors of dockless bike sharing demand. *IEEE Access,* 8, 66139–66149.

Martí, P., García-Mayor, C., & Serrano-Estrada, L. (2021). Taking the urban tourist activity pulse through digital footprints. *Current Issues in Tourism,* 24(2), 157–176.

McKercher, B., & Lau, G. (2008). Movement patterns of tourists within a destination. *Tourism Geographies,* 10(3), 355–374.

Md Khairi, N., Ismail, H., & Syed Jaafar, S. (2019). Tourist behaviour through consumption in melaka world heritage site. *Current Issues in Tourism,* 22(5), 582–600.

Miah, S., Vu, H., Gammack, J., & McGrath, M. (2017). A big data analytics method for tourist behaviour analysis. *Information and Management,* 54(6), 771–785.

Mitra, S. (2020). A different perspective to measure tourism attractiveness. *Current Issues in Tourism,* 23(14), 1718–1722.

Naumov, V., & Banet, K. (2020). Estimating parameters of demand for trips by public bicycle system using gps data. In *Advances in intelligent systems and computing* (Vol. 1091). Retrieved from https://doi.org/10.1007/978-3-030-35543-2_17 (AISC.)

Orellana, D., Bregt, A. K., Ligtenberg, A., & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. *Tourism Management,* 33(3), 672-682. Retrieved from https://www.sciencedirect.com/science/article/pii/S026151771100152X

Salas-Olmedo, M. H., Moya-Gómez, B., García-Palomares, J. C., & Gutiérrez, J. (2018). Tourists’ digital footprint in cities: Comparing big data sources. *Tourism Management,* 66, 13–25.

Salon, D., Conway, M., Wang, K., & Roth, N. (2019). Heterogeneity in the relationship between biking and the built environment. *Journal of Transport and Land Use,* 12(1), 99–126.

Smallwood, C. B., Beckley, L. E., & Moore, S. A. (2012). An analysis of visitor movement patterns using travel networks in a large marine park, north-western australia. *Tourism Management,* 33(3), 517–528.

Stangl, B., Prayag, G., & Polster, L. (2020). Segmenting visitors’ motivation, price perceptions, willingness to pay and price sensitivity in a collaborative destination marketing effort. *Current Issues in Tourism,* 23(21), 2666–2682.

Thees, H., Zacher, D., & Eckert, C. (2020). Work, life and leisure in an urban ecosystem - co-creating munich as an entrepreneurial destination. *Journal of Hospitality and Tourism Management,* 44, 171-183. Retrieved from https://www.sciencedirect.com/science/article/pii/S1447677020301637

Thimm, T., & Seebold, R. (2016). Past, present and future of tourist tracking. *Journal of Tourism Futures,* 2(1), 43–55.

Tracz, M., & Semczuk, M. (2018). Wpływ turystyki na zmianę funkcji przestrzennej, w przykładzie Krakowa. *Biuletyn Komitetu Przestrzennego Zagospodarowania Kraju PAN,* 272, 272–284.

Tran, T., Ovtracht, N., & D’Arcier, B. (2015). Modeling bike sharing system using built environment factors. *Procedia CIRP,* 30, 293–298.
Van der Ark, L. A., & Richards, G. (2006). Attractiveness of cultural activities in European cities: A latent class approach. *Tourism Management, 27*(6), 1408–1413.

Vogel, P., Greiser, T., & Mattfeld, D. (2011). Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia - Social and Behavioral Sciences, 20*, 514–523.

Wang, J., & Lindsey, G. (2019). Neighborhood socio-demographic characteristics and bike share member patterns of use. *Journal of Transport Geography, 79*, 102475.

Wang, K., & Akar, G. (2019). Gender gap generators for bike share ridership: Evidence from citi bike system in new york city. *Journal of Transport Geography, 76*, 1–9.

Zakrisson, I., & Zillinger, M. (2012). Emotions in motion: Tourist experiences in time and space. *Current Issues in Tourism, 15*(6), 505–523.

Zheng, W., Huang, X., & Li, Y. (2017). Understanding the tourist mobility using gps: Where is the next place? *Tourism Management, 59*, 267–280.

Zhu, H. (2020). Zone division and extraction of historic area based on big data. *Current Issues in Tourism, 0*(0), 1–22.

Önder, I., Koeberitz, W., & Hubmann-Haidvogel, A. (2016). Tracing tourists by their digital footprints. *Journal of Travel Research, 55*(5), 566–573.