This is How We Roll: Cultural Traits in Check-in Behaviour

Andreas Komninos
Computer & Information Science
University of Strathclyde
Glasgow, UK
andreas.komninos@strath.ac.uk

Jeries Besharat, John Garofalakis
Computer Engineering & Informatics
University of Patras
Patras, Greece
[besarat, garofala]@ceid.upatras.gr

Dimitrios Michelakis
School of Geosciences
University of Edinburgh
Edinburgh, UK
dimmihel@gmail.com

***Abstract*—We describe our analysis of three months worth of collected human-physical environment interactions (foursquare check-in) data for two European cities: a medium-sized city in the Mediterranean and a larger city in North West Europe. Our analysis shows that though the data generated by the citizens is scarce (i.e. on average less than 15 users checked in at any time during the day), nevertheless they can closely reflect a city's dynamics. The paper also discusses the strong suggestion of emergent cultural traits from the data and its use to power new pervasive applications based on human-physical interaction.**

Keywords— Check-in data, human-physical interaction, urban social context

I. INTRODUCTION (*HEADING 1*)

The concept of venues in social networks has become so popular, that a social network based exclusively on venues (Foursquare) is actively used by several million users, while other networks (Facebook, Google+) integrate formal venue representations with their existing structure. A large volume of human-physical interaction data is thus being generated daily from location-aware social networks. Research such as [1] demonstrates that this data is generated not randomly but under a very specific context, as users will only generate such data for venues they consider important, interesting or indicative of a social identity and lifestyle choice. Access to this data opens up significant new opportunities for the provision of informed services to a range of interested parties. For example, instead of getting walking directions from A to B, why not give directions that reflect the most "interesting" path instead of the shortest one? Such data could be used to automatically compile personalized tour guides or itineraries for visitors of a city. A further point of interest is guidance and recommendation of urban areas to visit under a given temporal context, instead of guiding users to individual POIs. This concept promotes exploration and offers the ability for visitors to explore alternatives, instead of targeting single POIs that can lead to disappointment (e.g. a nice restaurant that you arrive only to find it is fully booked and there are no other restaurants nearby). It could also be used to dynamically organize the routing and timing of public transport or taxis in order to maximize their efficiency and benefit both users and operators. Other services such as the planning of police presence, timing of traffic lights, mobile & public displays for visitors and tourists (see Fig. 1) or adaptive ambient lighting could be driven by such data. However before any of these visions can materialize, the data itself must be examined to see how useful it might in fact be.

Fig. 1. Visualising human-physical interactions as heatmaps in ubiqutous and mobile displays (from on-going experiments)

It appears that most current studies on social network venue data reflect on mobility patterns or land use. In most cases, analyses reflect on data from single locations. In [2] 34 cities are examined and a global underlying common pattern in the mobility of humans is found. The authors claim this to be in line with recent work on urban dynamics and organization, where cities have been shown to be scaled versions of each other, despite their cultural and historical differences, though this is contrasted by [3], where three US cities are found to be distinctly different in terms of venue clustering. Other studies such as [4][5] were able to distinguish between the configuration of venues or social activity in different areas within the same city. The question remains however, whether social network data is actually a realistic portrayal of human activity or whether the data (particularly mobility patterns) just happens to look realistic. So far, in a previous study [6] we were the first to uncover possible validation of the reliability of check-in data, where we found evidence that check-in data correlated well with traffic volume and air pollution data, although this was only for a medium-size city and for a different time frame.

Our work aims to examine the usefulness of human-physical interaction data from social networks, particularly for areas (cities) where the concentration of active users is not large and data might not be as rich as in a large metropolis. A further target is to examine whether such data can accurately capture changes through temporal context and also portray local cultural idiosyncrasies. For this purpose, we chose to contrast and compare data from two cities that are geographically and culturally different. This paper presents the outcomes of our comparison of user-physical interactions.

II. Data Collection Methodology

A. The cities

Our work focuses on two cities at the extremities of Europe: a medium-sized city South-East city (Patras, Greece) of approximately 200,000 inhabitants, and the larger city in the North-West (Glasgow, UK) of approximately 600,000 inhabitants. While both cities are in Europe, they belong to countries with distinct cultural differences and varying urban dynamics, which we present as background information that helps interpret the results that emerge from our dataset. The city rhythm is adequately described by the working hours, which is summarized in Table 1.

In Patras (and most of Greece), the typical working week is distinctly different from the typical Western 9-5 working week. Public sector and private sector employees have different working hours, as do retail businesses, which are closed on Monday, Wednesday and Saturday afternoons and all day Sunday. Retail is also closed between 2.30 and 5.30pm on all other days. Though the working schedule for the workforce is widely fragmented, the typically busy days in the city are Tuesday, Thursday and Friday. In contrast, Glasgow exhibits a much more standardized working week, typical of Western countries. Most businesses and professionals operate on a 9-5 schedule between Monday and Friday. Shops are open between 9-6 on all weekdays and Saturdays apart from Thursday, which sees shops open until 8pm. Shops are also open on Sundays, between 12 and 5pm. Both cities accommodate a large percentage of students amongst their population, each accommodating three universities. In Patras, all campuses are outwith the city centre, though most students live in it. In Glasgow, two of the universities have campuses located within the city center, while the third is outwith the city center area.

TABLE I. Professional Sector Opening Hours in Patras and Glasgow

	Patras	Glasgow
Public sector	7am-3pm (Mon-Fri)	9am-5pm (Mon-Fri)
Banking sector	9am-3pm (Mon-Fri)	9am-5pm (Mon-Fri)
Private offices & businesses	9am-9pm (Mon-Fri) break 2pm-5.30pm	9am-5pm (Mon-Fri)
Shops	9am-9pm (Mon-Fri), 9am-2pm Sat closed 2pm-5.30pm closed all evening Mon & Wed closed Sunday	9am-6pm (Mon-Sat) open till 8pm Thu 12pm-5pm (Sun)

B. The collected dataset

In [6] we described our collection methodology for obtaining data directly from the Foursquare API, which ensures that unlike other studies, e.g. [2][5], we do not depend on data that is tweeted at the time of creation, thus able to get a more complete set. We collected data for 100 days between 14 November 2012 and 26 February 2013, by querying the API every 30 minutes and recording the time, current check-ins and total check-ins for each venue. Here, it's important to discuss the concept of a check-in. Foursquare keeps a user checked into a place for a maximum of 3 hours or until the user checks in to another venue. As such, our data does not show distinct check-ins, but rather how many people *appear* to be checked into a venue at any point in time.

In Glasgow we discovered 1195 distinct venues, of which a total of 1164 demonstrated any check-ins. In Patras, 484 distinct venues were discovered, with 391 exhibiting some check-in activity. The ratios of check-in reports to venues was 8213:1 for Patras and 1718:1 for Glasgow, showing that in Patras, people check into fewer venues but more often, while Glasgwegians check in less and to a wider range of venues. Surprised by this disparity and keeping in mind that we take a count of appearances twice per hour, we tried to verify this assumption by calculating an estimate of the number of active Foursquare users that appear to be checked-in at any given hour. This gave an average of 8.24 users in Glasgow (stdev=5.03, max=14.83, min=0.52), surprisingly low, compared to Patras that exhibits an average of 15.45 users at any given hour (stdev=8.30 max=27.59, min=2.49). We had not anticipated this outcome, considering the penetration of smartphones in Greece (~25%)[1], the local cost of 3G connections (made significantly costlier because of the economic crisis and cuts), the lack of widespread adoption of free Wi-Fi compared to a richer Western city like Glasgow. Both statistics are still quite low and show that usage of Foursquare is not widespread in either city. The generally low adoption of Foursquare usage worldwide (>20 million users but just 31% of mobile users active on social networks are on Foursquare[2]), seems to contribute to our low surprise in the fact that the number of daily check-ins is still quite low. Our check-in data can be thus considered as "scarce" with regard to how frequently they are observed. Nevertheless, we wanted to explore whether urban dynamics could be uncovered from such scarce data.

III. How Reliable is Scarce Check-in Data?

Though local expert knowledge is a good starting point, we wanted to ensure that analysis of data captured through Foursquare could be validated against an objective quantifiable baseline. We sought other datasets indicative of human activity in a urban environment and considered thus air

[1] 42 major countries ranked by smartphone penetration rates, *Wired.com* (Dec 2011) [accessed 18/3/2014]

[2] SimplyZesty.com: Social media by the numbers: the latest site figures (April 2012) [accessed 18/3/2014]

pollution data, which coincide well with known urban dynamics [7]. Aiming to verify that this was not an isolated finding, we repeated the process for both cities. Data was obtained from government repositories. We considered nitrogen oxides (NO and NO_2) as these are the two pollutants most closely related to traffic. For Patras, we analysed data from 2009, which at the time was the most recent data that was available. For Glasgow, we analysed data from 2012, which mapped directly to the check-in data collection period. The output of that analysis is shown in Fig. 2. As is evident from the analysis of this data, both cities exhibit a measurable and distinctly different rhythm that matches our local knowledge as described in the previous section. In Patras, distinct activity peaks are noted in the morning and afternoons, coinciding with shop opening hours. In contrast, Glasgow exhibits an almost constant level of pollutants (particularly NO, typically emitted more than NO_2 by cars). We note here that most days are similar in pollution levels, while Saturdays and, in particular, Sundays seem to be the less polluted days. Comparing the atmospheric data with the check-in data, we observe that peaks occur slightly later than pollution data in the morning and earlier in the evening, which is expectable (people converge into to the city first and then start "checking in" to places, in the evening check-in activity gives way to transport activity). With this in mind, we found that the check-in data follows the atmospheric pollution patterns quite closely. For Glasgow, we observe a significant correlation between Foursquare check-ins and NO ($R_{(24)}=0.514,\ p<0.01$) as well as check-ins and NO_2 ($R_{(24)}=0.623,\ p<0.01$). In Patras, a significant correlation exists between check-ins and NO_2 ($R_{(24)}=0.548,\ p<0.01$) but not with NO, unlike the timeframe examined in our previous study[6]. However, we do find a significant and high correlation between NO and NO_2 ($R_{(24)}=0.719,\ p<0.01$), hence it can be argued that in both cases it appears that check-ins capture the activity levels in both cities as indicated by internal road traffic.

IV. Quantifiable Patterns in Check-in Data

Foursquare uses a multi-level category system for venues. In Glasgow, we detected places belonging to nine categories, while in Patras just five were reported. Particularly interesting considering that Foursquare check-ins are often made to display behaviour deemed "interesting" to peers [2], is the normalized appearances chart shown in Fig. 3. It is clear that in Patras, citizens feel more interested in reporting hedonic activities that relate to socializing with others. In contrast, Glasgow seems to value personal experiences more. The Transport & Travel category, which tops the Glasgow appearances, includes stations, stops and transit services, but also hotels and lodgings. Indicatively, out of these, the top three are Glasgow's two railway stations, followed by a hotel (8981, 4014 and 608 check-ins). The second most popular category is Arts & Entertainment followed by Nightlife and Shopping. All of these (except Nightlife) have more than twice the ratings that these categories receive in Patras.

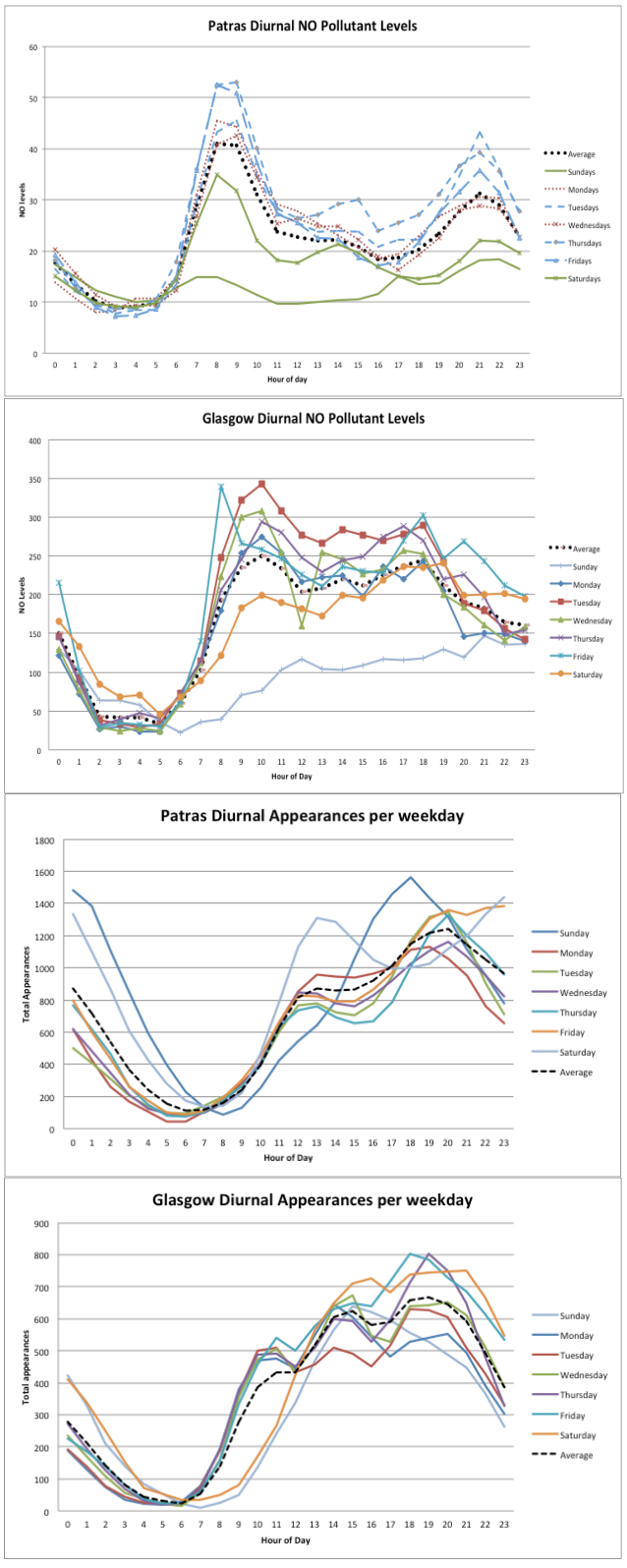


Fig. 2. Patras and Glasgow atmospheric pollution and check-in data diurnal breakdowns

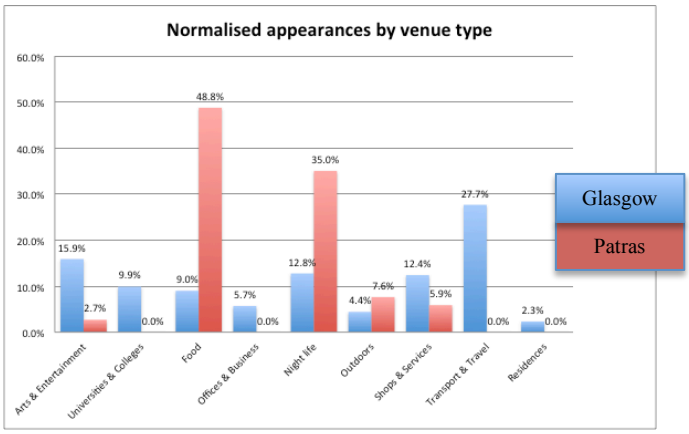


Fig. 3. Normalized venue category popularity

The data collection period included an interesting variation in normal patterns as it included the Christmas holiday season, celebrated in both countries. Fig 4. shows that Glasgow residents changed their behaviour significantly for travel (significant rise before Christmas and then rapid decline). Similarly, "Arts & Entertainment" venues seemed to gather fewer check ins over the holiday season, while for Patras there was little change (though again these venues take up a small proportion of check-ins).

An interesting comparison comes in change for "Shops & Services", rising in Glasgow after Boxing Day (Dec. 26), possibly because of the sales that start on that day. In contrast, we note that in Patras shopping activity reporting does not change (although this is generally not often reported, as seen in Figure 5). For Patras residents, we expected that the sales period would be a motivational factor to check in to shops more. Sales there began on January 15 but for that week, shopping check-ins are 0.4% less than the period average.

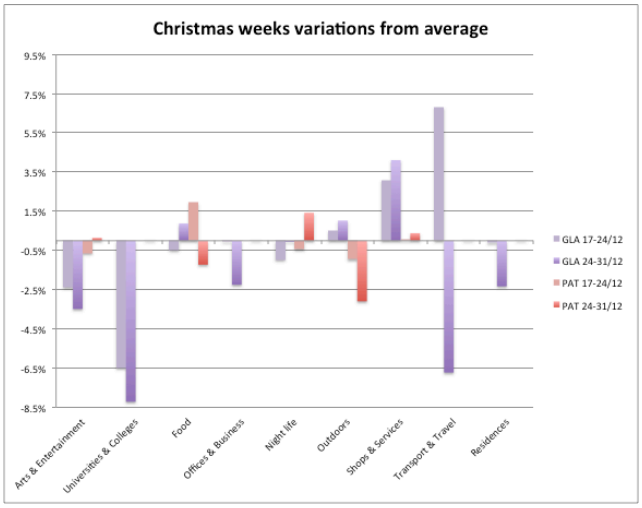


Fig. 4. Holiday season venue category popularity

To ensure that the differences in observed in the check-in behaviour are not a just a product of averaging out over the entire city, we sought to examine check-in behaviour in different parts of the city, defined not by abstract boundaries (e.g. neighbourhoods) but objectively defined classifications of areas. We hence used the total check-in data at the end of the period to break down the results of the historical popularity of venue types according to land use. We combined our point data with the GMES Urban Atlas[3], a pan-European land cover and use dataset for areas with over 100,000 inhabitants. Using Python and the Geospatial Data Abstraction Library (GDAL), we were able to identify the land use class that corresponds to each venue, by overlaying our point data in geographical space on polygons of land use types. Results indicate that irrespective of the land use type for both city centres (Continuous Urban Fabric, Discontinuous Dense Urban Fabric and Other Roads and Associated Land), Glasgow exhibits popularity at venues related to personal interest, whereas popular places in Patras are associated with social hedonism (e.g. Fig 5, 6 and 7).

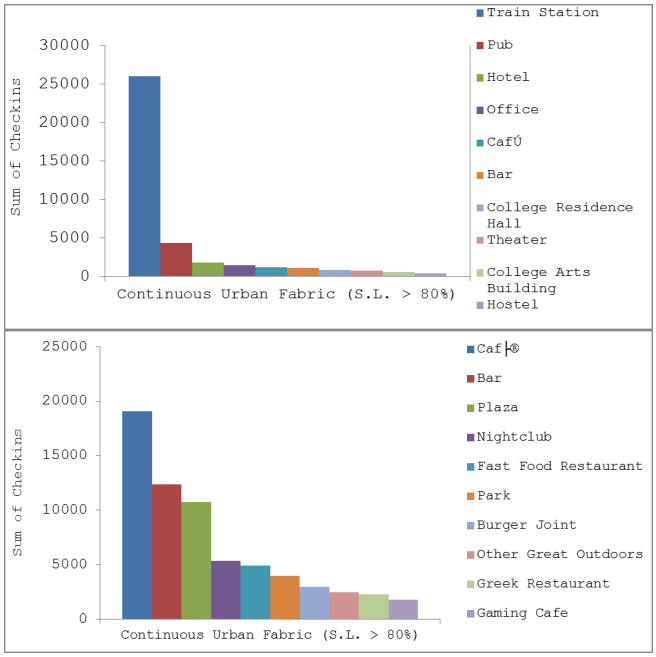


Fig. 5. Continuous urban fabric checkins for Glasgow (top) and Patras (bottom)

V. Do Cultural Traits Emerge In Check-In Data?

It could be said that check-in behaviour could be affected not only by attitudes toward use, but also by the availability of connected services in indoor and outdoor environments. We wondered thus if the image that emerges from our data is affected by the possibility that the low affordability of 3G in Greece compared to the UK, might have an impact on check-in patterns.

[3] European Environmental Agency GMES Urban Atlas http://www.eea.europa.eu/data-and-maps/data/urban-atlas/ [accessed 18/3/2014]

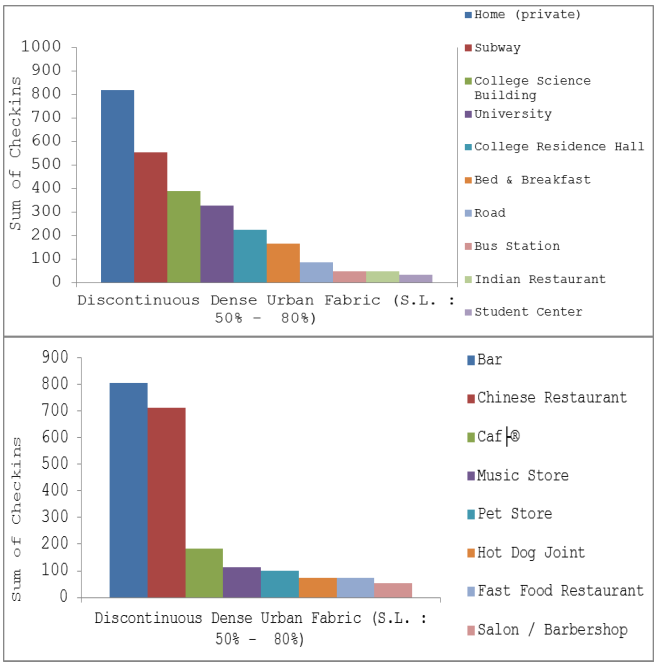


Fig. 6. Discontinuous dense urban fabric checkins for Glasgow (top) and Patras (bottom)

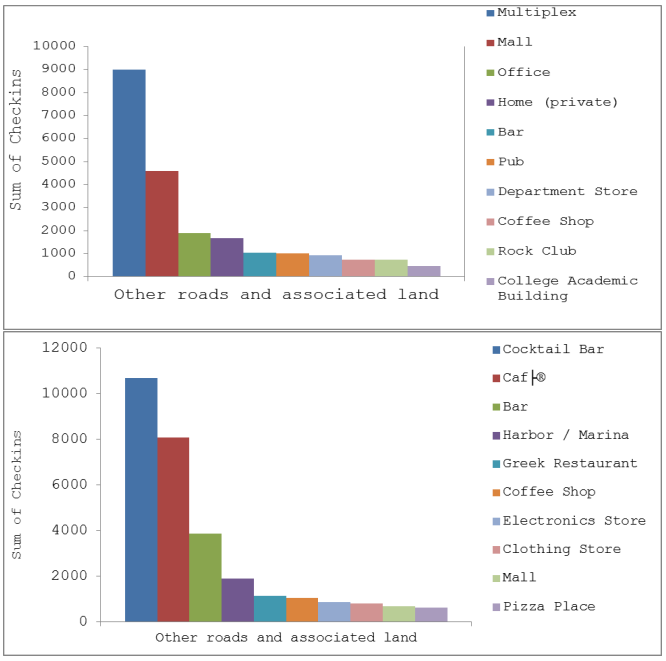


Fig. 7. Other roads and associalted lad checkins for Glasgow (top) and Patras (bottom)

However, the more frequent use and also the lack of check-ins in venue categories where Wi-Fi is likely to be available (e.g. homes or offices) in Patras, hinted that use is not just driven by the availability of access. To investigate further, we proceeded to harvest and mine Foursquare tips by extracting and mining Foursquare tips for venues, in order to separate out the venues for which users have reported that Wi-Fi is available (users often leave Wi-Fi passwords as venue tips to help other users). By doing so, we find no correlation between the reported existence of free Wi-Fi and the number of total check-ins in venues in Patras. Taking a sample of the 50 most popular venues with Wi-Fi and the 50 most popular without, we found a mean difference of $\Delta m=317.9$ check-ins in favour of those where free Wi-Fi is reported, though statistical significance is not found ($p=0.31$). These results indicate that check-in behaviour is more likely a product of cultural identity and not heavily influenced by ease of access.

Our focus thus shifted towards explaining the behavioural motivators behind check-in behaviour. The semantics of check-in actions are a recent research topic. Lindqvist et al. [1] discovered that users do not check in to places that they consider embarrassing (e.g. fast-food restaurants), interesting or visited frequently. Patil et al. [8] found check-ins to be driven strongly by projections of personal taste and image, such as a desire to indicate that a user likes a place, wanting to appear cool and interesting. Other reasons such as financial incentives (coupons) or promoting events did not motivate sharing as much. Similar findings were also discovered by Cramer et al. [9] who report the emergence of "*social and identity-driven uses such as sharing lifestyle, events and sharing of information that is interesting and enhances self-presentation*".

These observations bring to mind Hofstede's cultural value dimensions [10]. In charting national cultural values, he proposed that "*The British are a highly individualistic and private people*", while "*Greece is a collectivist culture […] which means that in this country people from birth onwards are integrated into the strong, cohesive in-group*". Kalogeraki [12] later argued that the UK remained an individualistic culture due to the complexity of multiculturality, while Greeks are mainly occupied with activities associated with their in-groups. The same research finds that Greeks still view concepts like "progress" as collective ideas and not as individual constructs – hence the modern trend of "checking in" might follow this behavioural aspect. Perhaps thus online check-in behaviour reflects real-world perceptions (and cultural norms) of what important activities are. This assumption is further encouraged by research that finds cultural traits to emerge in the use of social networks [10].

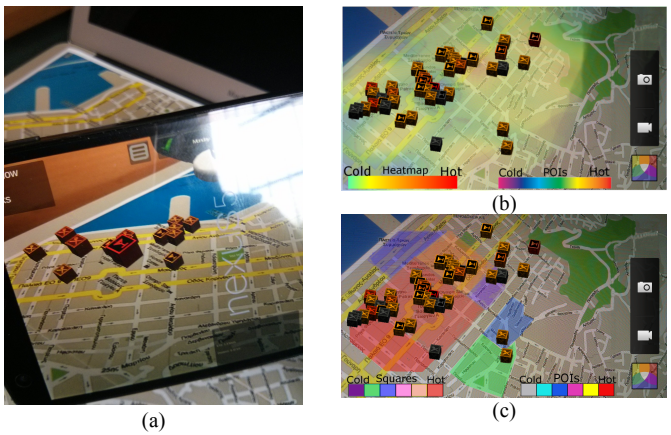
VI. Conclusions

To our knowledge, there are no studies that have, so far, managed to verify the degree to which check-in data correlates to an objective measure of human activity, which is independent of the accessibility of technology and network connectivity. Our paper's first contribution is that it finds evidence in two very different cities that scarce check-in data can be adequate to build an accurate picture of urban dynamics over time and that the data, though scarce, is reliable as it correlates well with independent human activity data sets (air pollution). Although a thorough analysis of data in more cities is needed to claim conclusive proof, we believe that this is a good first step that will convince future researchers to find innovative uses for this type of data.

We found strong suggestions that cultural differences in cities might also be accurately reflected by this data. Our opinion that check-in behaviour is driven mostly by cultural traits instead of technological factors, such as the availability of a cheap or free Internet connection, is backed by a lack of correlation between venue popularity and the availability of free wireless connectivity at these venues. Hence, our work offers quantitative results that back up qualitative studies such as [1, 8, 9]. Such studies have focused on single locations, hence missing the aspect of investigation of the role that cultural traits may play in check-in behaviour. We identify thus a need for further studies that will delve into the qualitative aspects of checking-in behaviour, to verify the emergence of these cultural traits.

Finally, one strong advantage of our work compared to previous studies (e.g. [2, 5] is that we are able to observe check-ins directly from Foursquare and not indirectly by mining Twitter feeds from users who have linked their Foursquare and Twitter accounts. Hence we feel our results are more trustworthy and offer a more complete picture of check-in behaviour than those of previous studies.

Our aim is to now explore how the inferred knowledge can affect citizens' interactions with the city, when presented to the users in a ubiquitous manner, such as desktop, mobile and AR applications, or through ubiquitous urban displays. We believe that such information can be invaluable to tourists and visitors of urban areas, helping them acquire better understandings of the popularities of entire areas and not just POIs. To this effect, we implemented an application running on a pervasive display and evaluated its effectiveness in [13], where it was found that the visualization of urban social context using a heatmap was received well by the tourists and helped them explore areas of the city, which they would have otherwise missed.



(a) (b) (c)

Fig. 8. Using Augmented Reality to place social context aware annotations on a traditional paper map. The size and colour of the POI markers is determined by spatiotemporally sensitive urban context information (a). The overall area social context is shown via heatmap (b) or in discrete areas (c). These are automatically generated from OpenStreetMap data and coloured according to Foursquare data for venues in the area.

Our work continues with an aim to provide both mobile applications that allow users to explore the social context of a urban environment by controlling both temporal and spatial parameters on a map (Fig. 1) as well as augmenting traditional paper maps with context-aware symbology, using a magic-lens augmented reality mobile application (Fig. 8a). The application shown has been built using the Qualcomm Vuforia SDK on an Android platform. Further, we are attempting to examine the correlation of check-in data with other types of human-physical interaction, such as tags, likes and tips, which we have been collecting as well. A final aspect of our ongoing work in the exploitation of urban social context is its representation for areas, through the use of alternative representations in Augmented Reality (Fig. 8b,c).

REFERENCES

[1] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman, "I'M the Mayor of My House: Examining Why People Use Foursquare - a Social-driven Location Sharing Application," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2011, pp. 2409–2418

[2] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo, "A Tale of Many Cities: Universal Patterns in Human Urban Mobility," *PLoS ONE*, vol. 7, no. 5, p. e37027, May 2012.

[3] J. Benner and C. Robles, Trending on Foursquare: Examining the Location and Categories of Venues that Trend in Three Cities. *Workshop on GIScience in the Big Data Age, GIScience2012*, 2012, pp. 27-34

[4] R. Lee, S. Wakamiya, and K. Sumiya, "Urban area characterization based on crowd behavioral lifelogs over Twitter," *Pers Ubiquit Comput*, vol. 17, no. 4, pp. 605–620, Apr. 2013.

[5] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh, The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. ICWSM'12. A. Vasalou, A. N. Joinson, and D. Courvoisier, "Cultural differences, experience with social networks and the nature of 'true commitment' in Facebook," *International Journal of Human-Computer Studies*, vol. 68, no. 10, pp. 719–728, Oct. 2010.

[6] A. Komninos, V. Stefanis, A. Plessas, and J. Besharat, "Capturing Urban Dynamics with Scarce Check-In Data," *Pervasive Computing, IEEE* , vol.12, no.4, pp.20,28, Oct.-Dec. 2013

[7] H. Chen, A. Namdeo, and M. Bell, "Classification of road traffic and roadside pollution concentrations for assessment of personal exposure," *Environmental Modelling & Software*, vol. 23, no. 3, pp. 282–287, Mar. 2008.

[8] S. Patil, G. Norcie, A. Kapadia, and A. Lee, "'Check out Where I Am!': Location-sharing Motivations, Preferences, and Practices," in *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 2012, pp. 1997–2002.

[9] H. Cramer, M. Rost, and L. E. Holmquist, "Performing a Check-in: Emerging Practices, Norms and 'Conflicts' in Location-sharing Using Foursquare," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, New York, NY, USA, 2011, pp. 57–66.

[10] G. Hofstede, *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, 2nd. ed. London: Sage, 2001.

[11] A. Vasalou, A.N. Joinson and D. Courvoisier, "Cultural differences, experience with social networks and the nature of "true commitment" in Facebook". *Intl. journal of human-computer studies* vol. 68, no.10, 2010, pp. 719-728

[12] S. Kalogeraki, "The Divergence Hypothesis in Modernization Theory Across Three European Countries: the UK, Sweden and Greece", *Culture Unbound* vol. 1, 2009, pp. 161-178

[13] A. Komninos, J. Besharat, D. Ferreira and J. Garofalakis, "HotCity: enhancing ubiquitous maps with social context heatmaps", *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia,* Dec. 2013, pp. 52-60