2012 IEEE International Conference on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing

Visualizing the Invisible Image of Cities

Thiago H. Silva*, Pedro O. S. Vaz de Melo*, Jussara M. Almeida*, Juliana Salles†, Antonio A. F. Loureiro*
*Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
†Microsoft Research, Redmond, WA, USA
{thiagohs, olmo, jussara, loureiro}@dcc.ufmg.br, jsalles@microsoft.com

***Abstract*—With the recent advances in technology, we have the opportunity to study the city dynamics at a large scale. A fundamental step is to be able to sense extensive areas. Participatory sensor networks have the potential to become a very fundamental tool to study social behavior at a large scale. Currently, there are some location sharing services, such as Foursquare, that record the user's location along the time, which is clearly one of dimensions in the city dynamics. In this work, we propose a technique called city image and show its applicability taking as examples eight different cities. The resulting image is a way of summarizing the city dynamics based on transition graphs, which map the movements of individuals in a PSN. This technique is a promising way to better understand the city dynamics, helping us to visualize the invisible images of cities.**

I. Introduction

With the advances in information and communications technology (ICT), we have the opportunity to study social behavior at a large scale, mainly in urban areas [1]. Cities are not identical and they evolve over time, presenting particular dynamics whose understanding is not a trivial task. A key step is to be able to sense vast zones. However, sensing such extensive regions poses many challenges such as the high cost of deploying an ICT solution at such large extent.

Smart phones [2] are taking center stage as the most widely adopted and ubiquitous computing device [3]. Those devices have enormous potential for the study of human social networks and human behavior "in vivo", in a natural context outside laboratories. Besides their computing power, smart phones are currently available with an increasing rich set of embedded sensors, such as GPS, accelerometer, microphone, camera, and gyroscope [3]. Sensing vast areas becomes more feasible when people carrying their portable devices (e.g., smart phones) collect data and collaborate among themselves. Systems that enable sensed data in this way are named participatory sensing systems (PSSs) [4]. In those networks, the shared data is not limited to sensor readings passively generated by the device, but also includes proactive user observations. There are many PSSs already deployed [5]. For instance, location sharing services that can be seen as location categorizing applications, such as Gowalla[1] and Foursquare[2]. This kind of application allows users to share their actual location associated with a specific place category (e.g., restaurant).

Location sharing systems are becoming very popular. Foursquare, created in 2009, registered 5 million users in December 2010, 10 million users in June 2011, and 20 million users in April 2012. With those datasets, we have an unprecedented opportunity to measure large scale city dynamics.

This work uses the concept of Participatory Sensor Network (PSN), a network derived from participatory sensing systems [6]. In this network, nodes are autonomous mobile entities (users) and the sensing activity depends on whether they want to participate in the sensing process [7]. In this work we show how a particular PSN (Foursquare in our case) can allow us to better understand the city dynamics, presenting a visualization technique denominated city image. The resulting image is a way of summarizing the city dynamics based on transition graphs, which map the movements among different categories of locations. As a case study, we apply this idea to eight different cities. We show that the city image is a powerful way to visualize the invisible image of cities. This technique might be useful in many cases. For instance, to help city planners to better understand the actual dynamics of a city, individuals of different categories (e.g, tourists) to decide where to go, taxi drivers (and other classes of workers) to meet individuals' demands, IT designers to propose customized software services for different categories of individuals, and for social behavior study.

The rest of the work is organized as follows. Section II presents the related work. Section III briefly discusses PSN concept. Based on one Foursquare dataset, Section IV presents fundamental characteristics used in the proposed case study. Section V details the city image technique, and demonstrates its usefulness considering different cities. Finally, Section VI presents the concluding remarks and future work.

II. Related work

Several previous studies have focused on the spatial properties of data shared in location sharing services such as Foursquare. For example, Cheng et al. [8] analyzed 22 million check-ins posted from more than 1,200 applications (Foursquare is responsible for 53.5% of the total). They found that users follow simple and reproducible patterns, and also that social status, in addition to geographic and economic factors, is coupled with mobility. Scellato et al. [9] present a study of the spatial properties of the social networks arising among users of Foursquare, Gowalla, and Brightkite. Among the results, the authors showed that 40% of social links happens below [100]km.

Cho et al. [10] investigated patterns of human mobility in three datasets, including check-ins of location sharing services and cellphone location data. They were particularly interested

[1]http://www.gowalla.com
[2]http://www.foursquare.com/

978-0-7695-4865-4/12 $26.00 © 2012 IEEE
DOI 10.1109/GreenCom.2012.62

in determining how often users move and where they go to, as well as how social ties may impact such movements. They observed that short-ranged travel is periodic both spatially and temporally and not affected by the social network structure, while long-distance travel is more influenced by social network ties. Noulas et al. [11] analyzed user check-in dynamics, showing that the distribution of number of check-ins at venues has heavy tail and power-law behavior. This is consistent with our findings [7]. They also found the presence of spatio-temporal patterns in Foursquare, showing considerable distinct patterns between weekdays and weekends (this result is also consistent with ours). However, this previous work, as the other aforementioned prior efforts, are focused mainly at investigating user mobility patterns, or social network properties and their implications. Next, we present other studies closer related to our proposal.

Doytsher et al. [12] analyzed a PSN from a different perspective. They presented an application that handles a social-spatial network (SSN). An SSN is a graph that consists of a social network, a spatial network, and life patterns that connect users of the social network to locations [12]. In the SSN-based application presented in [12], users can create queries such as "friends of Marge who buy at the same grocery store that she does", using a new query language.

The following proposals are concerned with a better understanding of the cities' invisible images, i.e., the dynamics of cities, including their particularities, inhabitants routines, etc. Cranshaw et al. [13] presented a model to extract distinct regions of a city that reflect current collective activity patterns. The idea is to expose the dynamic nature of local urban areas considering spatial proximity (derived from geographic coordinates) and social proximity (derived from the check-in distribution) of venues. Similarly, Noulas et al. [14] proposed an approach based on spectral clustering algorithm [15] to classify areas and users of a city by using venues' categories of Foursquare. The same group presented in [11] an analysis of place transitions in terms of check-in frequency. We also present in this work a similar evaluation, but differently from theirs, ours consider particular regions (median to large cities).

Lathia et al. [16] used a dataset from public transport usage to show that urban mobility is a viable way to better understand dynamics of a urban life. In their study they correlated the mobility of Londoners using the public transport system with the census-based indexes of the well-being from London's areas. The authors found interesting results, among them, socially-deprived communities in London tend to be visited more than wealthy ones.

Froehlich et al. [17] used a shared bicycling system dataset to gain insights into city dynamics and aggregated human behavior. Using that dataset, they showed that is possible to understand not just patterns of bicycling usage, but also underlying temporal and spatial dynamics of a city. They also demonstrated that simple predictive models are able to predict bicycle station usage with high accuracy.

Santi and Oliver [18] performed a social activity analysis in London, Paris, and New York using data collected from Foursquare. Among their findings, based on the information of the venue's category, they found that places from Food and Nightlife categories are the strongest social hub across the three cities. Our work differs from their work, because we study the city dynamics through people habits and routine, in a technique able to easy the city dynamics visualization.

III. PSN Derived from Location Sharing Services

Participatory sensing is the process where individuals use mobile devices and cloud services to share sensed data [4]. Usually participatory sensing systems consider that the shared data is generated automatically/passively, by sensor readings from the device, but we also consider manually/proactively, user-generated observations. Participatory sensing with this characteristic has been called ubiquitous crowdsourcing [19].

Location sharing services, such as Gowalla and Foursquare, are also examples of participatory sensing applications. The sensed data is an observation (check-in) of a particular place that indicates, for instance, a restaurant in a specific place. By analyzing a dataset from this service, one is able to discover what is around you, or receive recommendations of places to visit. In the remainder of the paper we will use the word "check-in" to refer to an event when time and location of a particular user is recorded or, in a PSN context, sensed.

From a participatory sensing system we can derive a PSN [7], where the user's portable device is the fundamental building block. Individuals carrying these devices are able to sense the environment and to make relevant observations at a personal level. Thus, each node in a PSN consists of the user plus his mobile device. As in WSNs, the sensed data is sent to a server, or the "sink node". But differently from a WSN, PSNs have the following characteristics:

- nodes are autonomous mobile entities (users);
- the cost of the network is distributed among the nodes, making the network globally scalable;
- sensing depends on the nodes that will participate in the sensing process;
- nodes transmit the sensed data directly to the sink;
- sink only receives the data and does not have direct control over the nodes and
- nodes do not face severe energy constraints.

In Figure 1, we show an example of a PSN comprised of location sharing services, which we analyze in the following sections. This figure represents four users at three different points in time (Figures 1a, 1b and 1c). Locations shared by users at each time are pointed with dashed arrows. Note that users do not necessarily participate in the system all the time. After a given time, we can analyze this data in very different ways. For instance, in Figure 1d, we show a graph where nodes represent shared locations and edges connect shared locations by the same user. With this graph we can extract many valuable information, such as regular trajectories by a user. Moreover, given the ubiquity of smart phones, it is possible to include people with different interests from different parts of the world, providing a remarkably globally scalable and affordable infrastructure, as we show in Figure 2.

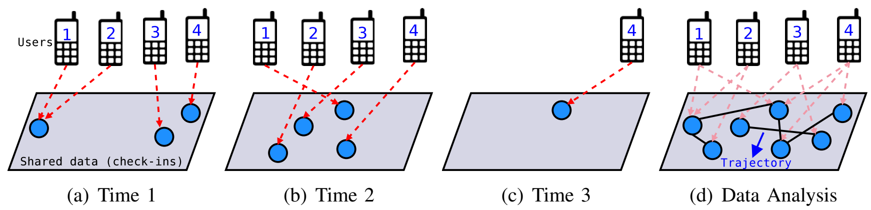


Fig. 1. PSN analyzed: location sharing services (Foursquare)

IV. PSN Characteristics

Here we present some characteristics of a participatory sensor network (PSN) derived from Foursquare, a popular location sharing service. The idea is not to present all characteristics, but only give an idea about the potential of this sort of network (Section IV-B) and justify some decisions in the results presentation (Section IV-C). More characteristics and challenges of PSNs can be found in our previous work [7].

A. Data Description

Foursquare is a location sharing service (also known as location-based social network) created in 2009 and is one of the most popular systems of its category, registering more than 20 million users in April 2012. A Foursquare application is designed for mobile devices, allowing users to share their location to their friends (this procedure is called check-in)

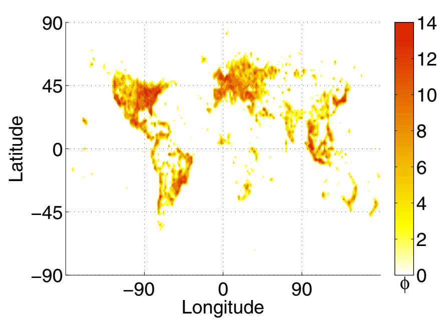


Fig. 2. All sensed locations. The number of locations n per pixel is given by the value of ϕ displayed in the colormap, where $n = 2^{\phi} - 1$.

We collected a dataset of Foursquare directly from Twitter[3], since Foursquare check-ins are not publicly available, by default. Approximately 4.7 million tweets containing check-ins were extracted from Twitter, each one providing a URL to the Foursquare website, where information about the geographic location of the venue was acquired. This dataset contains for each check-in the latitude, longitude, venue's id and time, and venue's category (essential for the analysis presented in Section V). In Table I, we describe these categories. In this work, we consider that Home and Work make more sense being distinct categories. To do that we associated sub-categories related to home, as a new category named Home. We did the same for the new Office (abbreviation "Offi") category. This dataset represents one week of April 2012, and contains 4,672,841 check-ins, and 1,929,237 venues.

[3]http://www.twitter.com

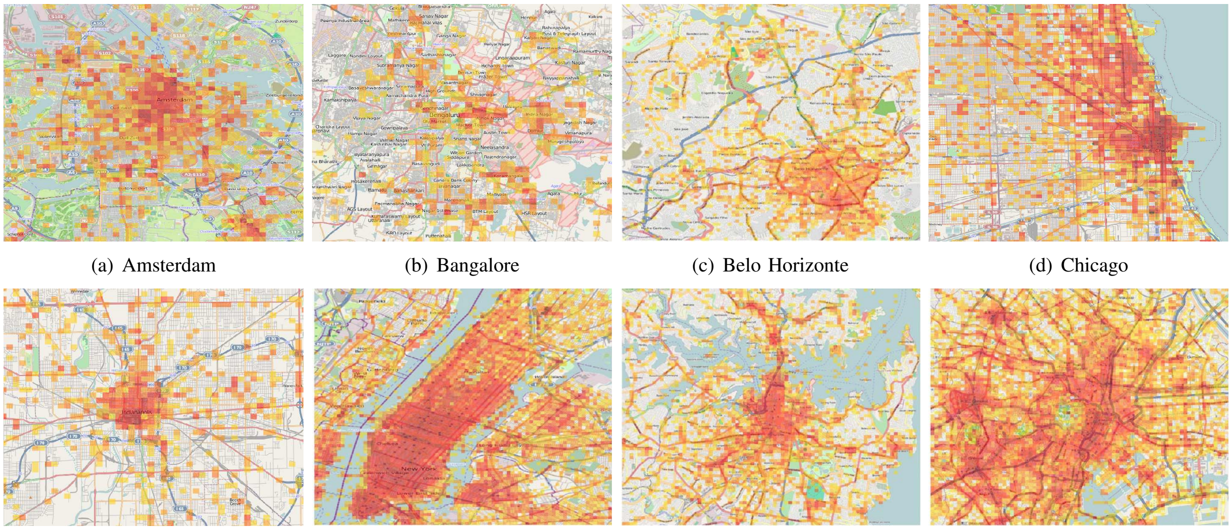
TABLE I
Foursquare Categories

Name	Abbreviation	Sub-categories examples
Arts& Entertainment	A&E	Comedy Club, Movie Theater, Museum, Casino
College & Education	Edu	College Lab, Fraternity House, Student Center
Food	Food	Bakery, Restaurant, Coffee Shop, Pizza Place
Home, Work and Others	H/O	Home, Residential Building, Factory, Conference Room
Great Outdoors	Outd	Baseball Field, Surf Spot, Park, Cemetery
Nightlife Spot	NL	Bar, Rock Club, Nightclub, Strip Club
Shop & Service	Shop	Shoe Store, Nail Salon, Deli or Bodega, Music Store
Travel Spot	Trvl	Airport, Subway, Embassy or Consulate, Hotel

B. Network Coverage

Figure 2 depicts the coverage in the PSN formed from our dataset, which can be very comprehensive in a planetary scale. Despite that the sensing activity in some continents, such as North America and Europe, are higher than in others, such as Oceania and Africa, we still have a global magnitude of coverage. With a PSN we can achieve such scale of sensing at very small cost.

We now analyze the coverage in eight cities in different continents: Amsterdam/Netherlands (7457 check-ins), Bangalore/India (1663 check-ins), Belo Horizonte/Brazil (13571 check-ins), Chicago/USA (28366 check-ins), Indianapolis/USA (6535 check-ins), New York/USA (70705 check-ins), Sydney/Australia (4346 check-ins), and Tokyo/Japan (85364 check-ins). Figure 3 shows, for each city, the heatmap of the sensing activity in these cities, enabling the visualization of the sensing activity. In the heatmap, the darker the color, the higher the number of check-ins in that area. We can see that in some cities the coverage is high, such as Chicago (Figure 3c), New York (Figure 3f), and Tokyo (Figure 3h). On the other hand, there are some cities with low sensing coverage, such as Bangalore (Figure 3h). Many factors influence the sensing coverage, like number of inhabitants, geographic aspects, cultural aspects, just to mention a few.



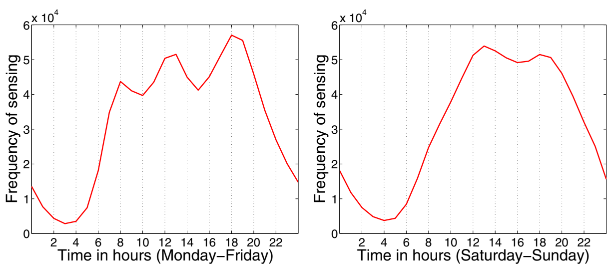
(a) Amsterdam (b) Bangalore (c) Belo Horizonte (d) Chicago

(e) Indianapolis (f) New York (g) Sydney (h) Tokyo

Fig. 3. Sensed locations in eight cities. The number of check-ins in each area is represented by a heatmap. The color varies from yellow to red (high intensity).

C. Seasonality

We now analyze how the seasonal behavior of humans affects the data sharing. We consider the location sharing pattern for the analyzed dataset[4]. Figure 4a shows the average number of check-ins during each hour from Monday to Friday. Figure 4b shows the same information for Saturday and Sunday. As expected, the network actuation presents a diurnal pattern, meaning that during the dawn the sensing activity is very low. Considering weekdays (Figure 4a), it is also possible to observe three peaks during the day, around [8:00]am (breakfast), [1:00]pm (lunch), and [6:00]pm (dinner). On weekends, there is no peak activity in the morning, the lunch peak happens around [1:00]pm, and the dinner peak is almost flat (from [6:00]pm to [7:00]pm). We can also observe that the activity is more intense on weekends. This indicates that there are four classes of behavior: day and night, for weekdays and weekends.

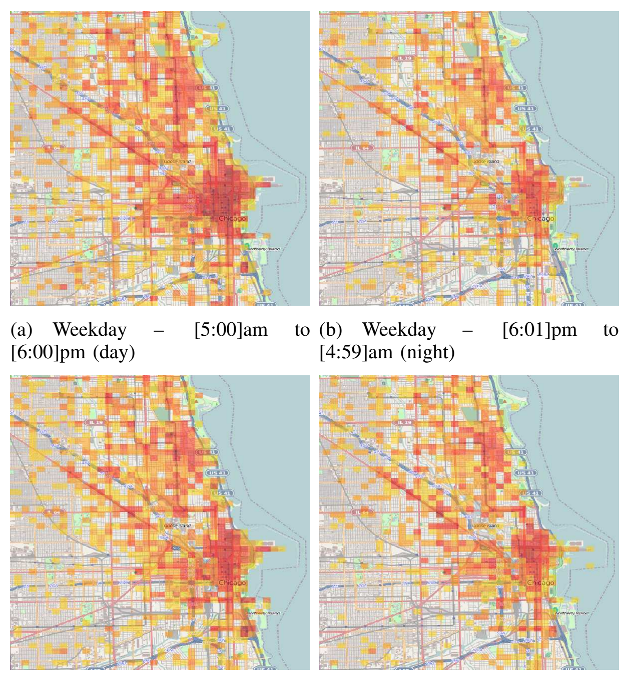


(a) Monday to Friday (b) Saturday and Sunday

Fig. 4. Weekdays and weekend location sharing patterns

These differences in classes can be visually detected. Figure 5 shows, for Chicago, the heatmap of its check-in activity. In the heatmap, the darker the color, the higher the number of check-ins in that area. As we can see, there are four classes of behavior. For this reason, we consider them in Section V, where we present the analysis for our eight cities.

[4]Timestamps were normalized according to the local time of the check-in



(a) Weekday – [5:00]am to [6:00]pm (day)
(b) Weekday – [6:01]pm to [4:59]am (night)
(c) Weekend – [5:00]am to [6:00]pm (day)
(d) Weekend – [6:01]pm to [4:59]am (night)

Fig. 5. All sensed locations in Chicago for four different periods. The number of check-ins in each area is represented by a heatmap. The color varies from yellow to red (high intensity).

V. Visualizing the Invisible Image of a City

As we mentioned before, the sensing activity in a PSN is performed by mobile individuals who decide to share their information. Different from traditional mobile Wireless Sensor Networks, the nodes in a PSN move accordingly to their routines or local preferences, and this generates the skewed behavior in the sensing activity that we have seen so far [7]. Besides this, if individuals have the incentive to check-in in every location he/she visits, then we could clearly observe the flow of people in a given city during a given time. In this direction, we present the construction of a transition graph, which maps the movements of individuals in a PSN.

Similarly to Kostakos et al. [20], we also believe that cities are not identical and also evolve over time. In this direction, we propose the city image, a square matrix that summarizes the city dynamics. This matrix is constructed from two transition graphs. First, we construct a transition graph $G(V, E)$, where the nodes $v_i \in V$ are the main categories of the locations and an edge (i, j) exists from node v_i to node v_j if at some point in time an individual performed a check-in in a location categorized by v_j just after performing a check-in in a location categorized by v_i. The weight $w(i, j)$ of an edge is the total number of transitions that occurred from node v_i to node v_j.

The first requirement for a transition to be characterized is that the check-ins must be performed consecutively and by the same individual. The second one is discard continuous check-ins at the same venue (the number of transitions not considered is not significant, the maximum percentage occurred in the Amsterdam dataset: 1.8%). Moreover, the transition must have occurred at the same "social day". We define a "social day" as the day which starts at [5:00]am instead of [12:00]am, since we are interested in capturing the nightlife transitions as well. If a transition occurred between two different "social days", we will only consider it if the time interval between them is lower than four hours. We tested several different policies for characterizing transitions and the results are very similar, since only a small percentage of transitions are discarded/considered when we vary the policy.

After constructing G, we create ten random graphs $G_{Ri}(V, E_{Ri})$, where $i = 1, \ldots, 10$ and each one is constructed using the same number of transitions used to construct G. However, instead of considering the actual transition $v_i \to v_j$ performed by an individual, e.g., "Smith", we randomly pick a location category to replace v_j, simulating, then, a random walk for this individual.

After that, we verify if the distribution for the randomly generated edge weight values for $G_{R1..10}$ follow a normal distribution $N(\overline{w}, \sigma_w)$. In an affirmative case, we compute, respectively, the mean $\overline{w}$ and the standard deviation σ_w for the edge weights in G_{Ri}. We define the *indifference range* as the interval $(\overline{w} - 3\sigma_w, \overline{w} + 3\sigma_w)$, which is expected to contain 99.73% of the randomly generated edge weight values, since the edge weights follow a normal distribution $N(\overline{w}, \sigma_w)$. Analogously, we define the *rejection range* as the interval $[-\infty, \overline{w} - 3\sigma_w]$ and the *favouring range* as the interval $[\overline{w} + 3\sigma_w, \infty]$. Otherwise, we calculate the maximum (max) and minimum (min) values of the randomly generated edge weight values for G_{Ri} (procedure was executed for just cases). We, then, define the *indifference range* as the interval (min, max). The *rejection range* is defined as the interval $[-\infty, min]$ and the *favouring range* as the interval $[max, \infty]$

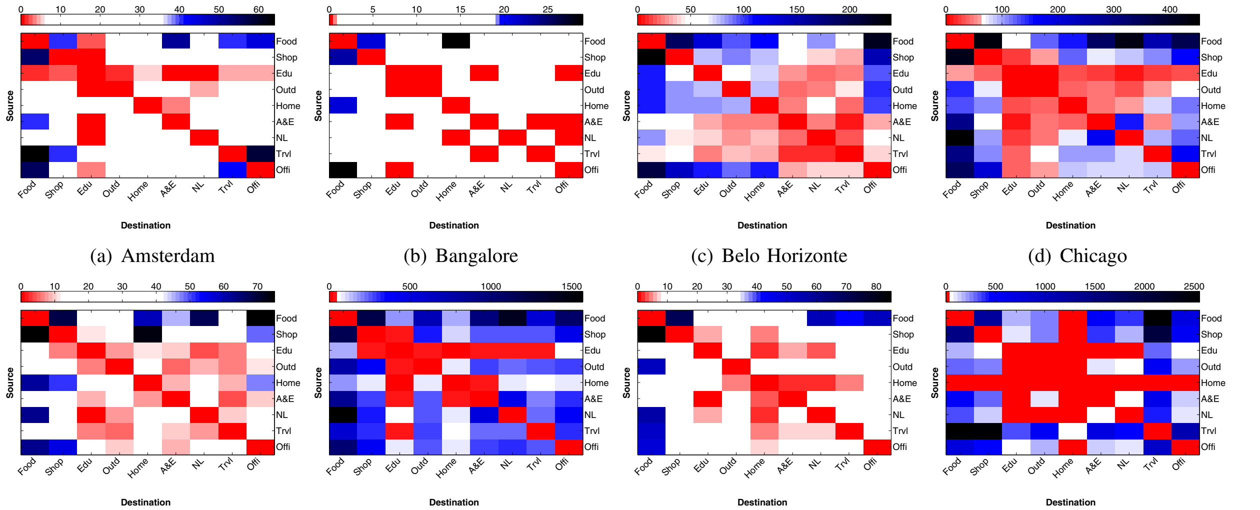
A. The DNA for our eight cities

Figure 6 presents the city image for all considered cities with no distinction of periods. In other words, the data used to generate the city image for each city is composed of all time periods, without any separation, such as weekdays and weekends. This information is useful because it represents a general picture of the city. The drawback is that we lose some important information without considering different periods apart. To illustrate that, compare the general city image for Amsterdam in Figure 6a, with the city image considering different periods, for the same city, in Figure 7 (the explanation for this figure is given below). First, we can easily visualize the distinction between them. Moreover, note that nightlife transitions are more favorable to happen at night, especially on weekends.

Figures 7–14 present the city image for the cities Amsterdam, Bangalore, Belo Horizonte, Chicago, Indianapolis, New York, Sydney, and Tokyo, respectively. Each figure is composed by four sub-figures, which represent four different periods: day, from [5:00]am to [6:00]pm, on weekday (sub-figure "a"); day on weekend (sub-figure "b"); night, from [6:01]pm to [4:59]am, on weekday (sub-figure "c"); and night on weekend (sub-figure "d"). We show on which interval the transitions between categories fall for those cities. Each cell represents the willingness of a transition from a category at a given place (vertical axis) to another category (horizontal axis). Red colors indicate a natural rejection for that transition to occur, i.e., the edge weight for this transition falls in the *rejection range*. Blue colors mark transitions that occur frequently, i.e., the edge weight for this transition falls in the *favouring range*. Finally, white color indicates indifference, i.e., the edge weight for this transition falls in the *indifference range*.

Observe for those figures that the city image captures the city dynamics in a very summarized way. Note how different are the dynamics of the cities and how they change from weekdays to weekends, and from day to night. Moreover, observe the main diagonal of the matrices, which indicates a tendency of not having consecutive check-ins at the same category. Also, it is easy to understand what are the most and least favored places and transitions of each city in a given day.

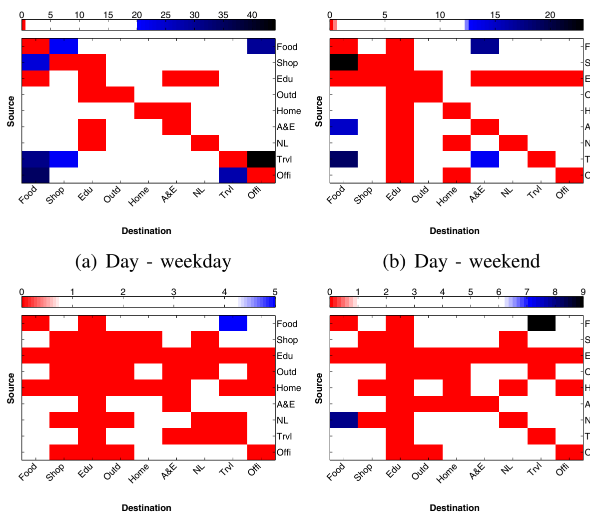
In general, using the city image it is possible to distinguish the routines of the inhabitants seen in two cities. For instance, in Bangalore (Figure 8) we observe the lack of favorable transitions considering the category `nightlife`, in all periods analyzed. On the other hand, this is strong favorable to happen in Chicago (Figure 10), and New York (Figure 12) not just on a weekend at night, as it is more common to happen, but also on weekdays at night. On weekends at night



(a) Amsterdam (b) Bangalore (c) Belo Horizonte (d) Chicago

(e) Indianapolis (f) New York (g) Sydney (h) Tokyo

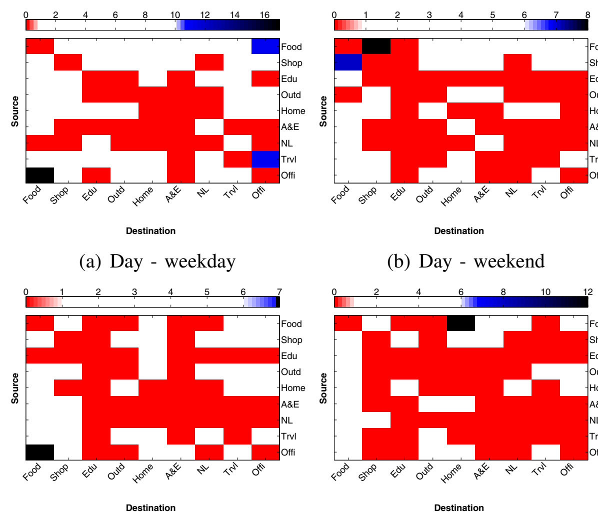
Fig. 6. The general city image, which does not consider different periods separately of all cities. Each cell represents the willingness of a transition from a given category at a given place (vertical axis) to another category (horizontal axis). Red colors represent rejection, blue colors represent favoring and white represents indifference.



(a) Day - weekday (b) Day - weekend

(c) Night - weekday (d) Night - weekend

Fig. 7. The image of Amsterdam for different periods.



(a) Day - weekday (b) Day - weekend

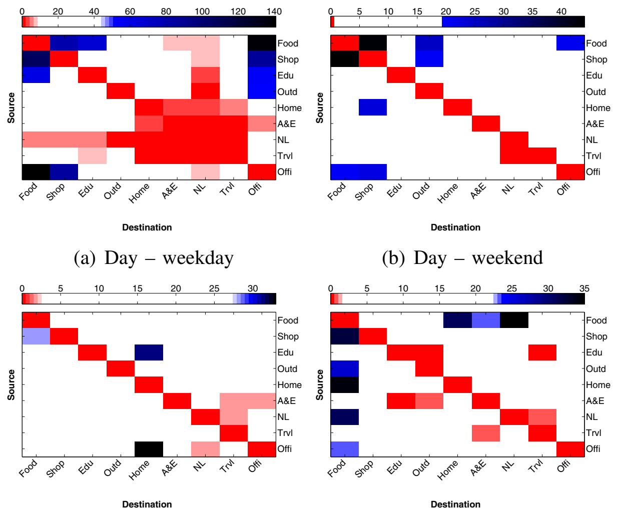
(c) Night - weekday (d) Night - weekend

Fig. 8. The image of Bangalore for different periods.

inhabitants from Bangalore are very favorable to perform the transitions: `food` → `home`. This might be explained by cultural differences existing among those cities.

Inhabitants from Belo Horizonte (Figure 9) are highly favorable to perform the transition containing the category `education`. This comes with no surprise since it is an important hub of education in Brazil. In the city image of this city is worth noting that the transition `education` → `office` is favorable. In Belo Horizonte is common to find students who have jobs (part-time of full-time) while they are still students. This explain the transition `education` → `home` on weekdays at night, because students who have a full-time job have the opportunity to go to school at night. Related to this discussion one unexpected result was the tendency of rejection of any transition involving the category `education` in Chicago (in any period analyzed), since it has been a world center of higher education and research with several universities.

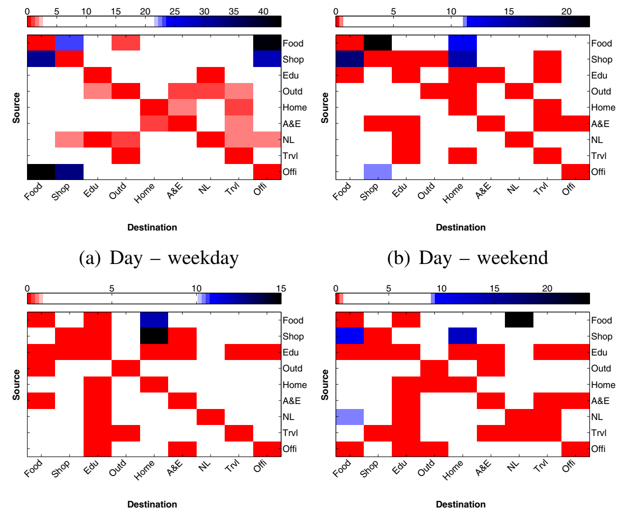
In Amsterdam (Figure 7), the most favored transition is from `travel` → `office` in weekdays during the day. Similar trends are common in many cities, for instance, in New York and Tokyo (Figure 14). While some cities, such as Indianapolis (Figure 11), Belo Horizonte, and Sydney (Figure 13), do not present favorable transitions containing the category `travel`



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

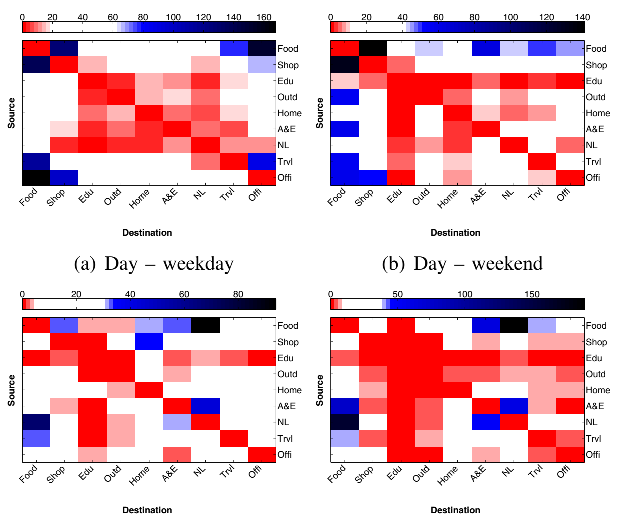
Fig. 9. The image of Belo Horizonte for different periods.



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

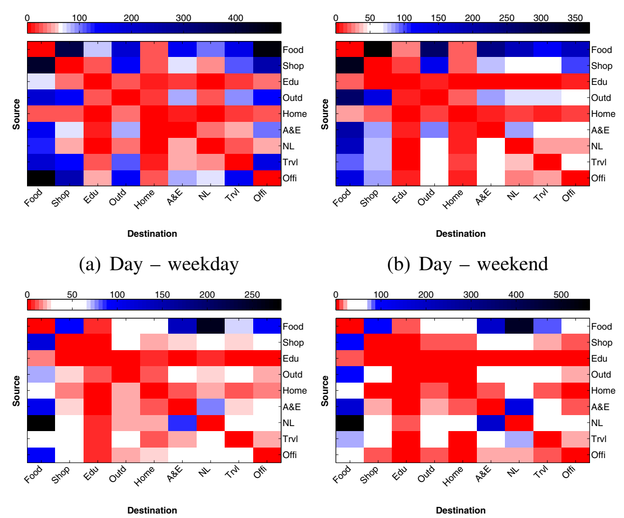
Fig. 11. The image of Indianapolis for different periods.



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

Fig. 10. The image of Chicago for different periods.



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

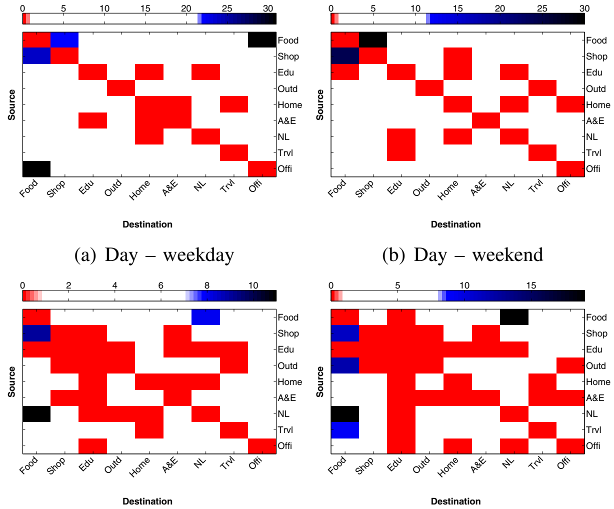
Fig. 12. The image of New York for different periods.

on weekdays during the day. This could be associated with a higher number of people that drive in order to get in their destination.

The city image technique, as illustrated above, is a interesting way to better understand the invisible image of a city. The usefulness for that are various, ranging from helping city planners to better understand the actual dynamics of a city, to providing tourists with one more type of information that might be useful, for example, in deciding which city to visit. The transition tendencies further serve as fundamental information for social behavior study. It is important to point out some limitations of our dataset. First, it reflects the behavior of part of the citizens. Second, since we just have a sample of the activity occurred in the year, bad weather conditions might have affected the number of check-ins in some places, especially those under outdoor category. This does not invalidate our technique, but prevent us to do certain kind of assertions.

VI. Conclusions and Future Work

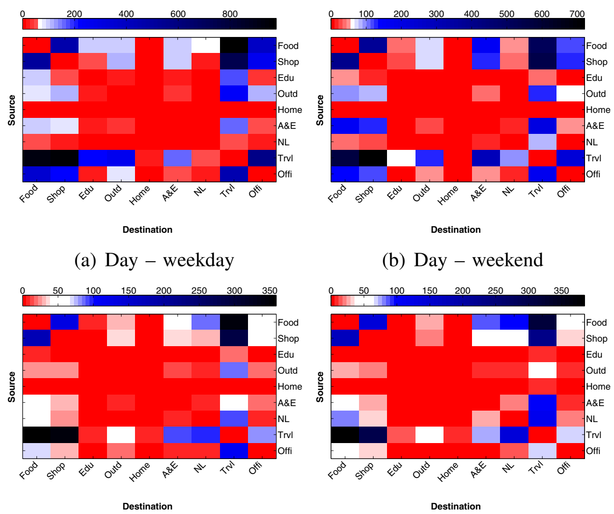
Participatory sensor networks have the potential to become a very fundamental tool to study social behavior at a large scale. A simple and important sensing data is the user's location, which clearly captures one of the dynamics dimension of a city. Currently, there are some location sharing services, such as Foursquare. This kind of application allows users to share their actual location associated with a specific place category. Based on those datasets, we have an unprecedented



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

Fig. 13. The image of Sydney for different periods.



(a) Day – weekday (b) Day – weekend

(c) Night – weekday (d) Night – weekend

Fig. 14. The image of Tokyo for different periods.

opportunity to measure large scale city dynamics.

In this work we showed the applicability of PSNs. Using a Foursquare dataset, we presented the applicability of this technique, called city image, taking as examples eight different cities. The resulting image is a way of summarizing the city dynamics based on transition graphs, which map the movements of individuals in a PSN. This technique is a promising way to better understand the city dynamics, helping us to visualize the invisible images of cities.

At this time, we are working in two main directions. First, we are analyzing other types of participatory sensing systems to complement our analysis on location sharing services. Second, we are working to enable the analysis of a large number of cities in the world using the city image technique.

ACKNOWLEDGMENT

This work is partially supported by the INCT-Web (MCT/CNPq grant 57.3871/2008-6), and by the authors' individual grants and scholarships from CNPq, CAPES (scholarship 7356/12-9), and FAPEMIG.

REFERENCES

[1] D. Lazer *et al.*, "Computational Social Science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[2] G. Miller, "The Smartphone Psychology Manifesto," *Perspectives on Psychological Science*, vol. 7, no. 3, pp. 221–237, 2012.

[3] N. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. Campbell, "A Survey of Mobile Phone Sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140 –150, sept. 2010.

[4] J. B. et al., "Participatory sensing," in *Workshop on World-Sensor-Web (WSW'06)*, 2006, pp. 117–134.

[5] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment Framework for Participatory Sensing Data Collections," in *Proc. of the 8th International Conference on Pervasive Computing (Pervasive'10)*, 2010, pp. 138–155.

[6] J. Goldman *et al.*, "Participatory Sensing: A Citizen-powered Approach to Illuminating the Patterns that Shape our World," Woodrow Wilson Center for International Scholars, White Paper, May 2009.

[7] T. H. Silva, P. O. S. Vaz de Melo, J. M. Almeida, and A. A. F. Loureiro, "Uncovering Properties in Participatory Sensor Networks," in *Proc. of the 4rd ACM International Workshop on Hot Topics in Planet-scale Measurement (HotPlanet'12)*, June 2012.

[8] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, "Exploring Millions of Footprints in Location Sharing Services," in *Proc. of the Fifth Int'l Conf. on Weblogs and Social Media (ICWSM'11)*, 2011.

[9] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo, "Socio-spatial Properties of Online Location-based Social Networks," in *Proc. of the Fifth Int'l Conf. on Weblogs and Social Media (ICWSM'11)*, 2011.

[10] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proc. of the 17th ACM Int'l Conf. on Knowledge Discovery and Data Mining (KDD'11)*, 2011, pp. 1082–1090.

[11] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "An Empirical Study of Geographic User Activity Patterns in Foursquare," in *Proc. of the Fifth Int'l Conf. on Weblogs and Social Media (ICWSM'11)*, 2011.

[12] Y. Doytsher, B. Galon, and Y. Kanza, "Querying Socio-spatial Networks on the World Wide Web," in *Proc. of the 21st Int'l Conf. Companion on World Wide Web (WWW'12)*, 2012, pp. 329–332.

[13] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh, "The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City," in *Proc. of the Sixth Int'l Conf. on Weblogs and Social Media*, 2012.

[14] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "Exploiting Semantic Annotations for Clustering Geographic Areas and Users in Location-based Social Networks," in *Proc. of the Fifth Int'l Conf. on Weblogs and Social Media (ICWSM'11)*, 2011.

[15] U. Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec. 2007.

[16] N. Lathia, D. Quercia, and J. Crowcroft, "The Hidden Image of the City: Sensing Community Well-Being from Urban Mobility," in *Proc. of 10th Int'l Conf. on Pervasive Computing*, 2012.

[17] J. Froehlich, J. Neumann, and N. Oliver, "Sensing and Predicting the Pulse of the City through Shared Bicycling," in *Proc. of the 21st Int'l Jont Conf. on Artifical intelligence (IJCAI'09)*, 2009, pp. 1420–1426.

[18] S. Phithakkitnukoon and P. Oliver, "Sensing Urban Social Geography Using Online Social Networking Data," in *Proc. of the Fifth Int'l Conf. on Weblogs and Social Media (ICWSM'11)*, 2011.

[19] A. J. Mashhadi and L. Capra, "Quality Control for Real-time Ubiquitous Crowdsourcing," in *Proc. of the 2nd Int'l Workshop on Ubiquitous Crowdsouring (UbiCrowd'11)*, 2011, pp. 5–8.

[20] V. Kostakos *et al.*, "Understanding and Measuring the Urban Pervasive infrastructure," *Personal and Ubiquitous Computing*, vol. 13, no. 5, pp. 355–364, Jun. 2009.