Evaluation of Geosocial Data Representativeness based on Human Spatial Memory Perspective: The example of Foursquare check-ins

Teriitutea Quesnot and Stéphane Roche

Université Laval, Center for Research in Geomatics
Quebec City (QC), Canada
teriitutea.quesnot.1@ulaval.ca
stephane.roche@scg.ulaval.ca

Abstract

Once combined with participatory mapping and mobile Internet, Geolocation made possible the emergence of a new type of a social network called Location-Based Social Network (LBSN). Contrary to the personal geolocated services, LBSN such as Foursquare allows its users to be permanently connected with other people (friends) in a specific social network and several kinds of spaces. Thus, the question is not so much "Where am I?" as "What is around me?" (services, people, traffic, disturbances, shops…), "What can I expect?" and "How do I get there?". By sharing their spatial location via a "check-in", users produce geo-traces of their daily routes and movements. These geolocated data are considered as a new kind of geospatial information whose potential of analysis has been shown by many researchers (e.g. Cranshaw *et al*, 2012; Stefanidis *et al*, 2011). To our knowledge, none of these has explored the representativeness of Foursquare check-ins from the human spatial memory perspective, which is in fact one specific part of spatial cognition (McNamara & Valiquette, 2004). Since our "mental map" tends to be topologic (Davies *et al*, 2010), we could argue that the LBSN users activity allows identification of significant landmarks (Lynch, 1970) – also described as "anchor points" by Golledge (1978) - and by extension, delimitation of cognitively relevant areas. The main objective of this Ph.D. research consist of evaluating these areas by making interviews of Quebec City's Foursquare Mayors in order to analyze distances and directions perceived between places whose cognitive significance will be assessed by Foursquare check-ins data spatial analysis. Indeed, several researches have demonstrated that distances tend to be underestimated from relevant landmarks (e.g. Sadalla *et al*, 1980) and direction estimates appear to be more accurate (Golledge, 1999). This research aims at providing evidences of the reliability of this specific kind of Geodata to generate real-time "smart landmarks" for the future smart cities. Therefore, we could seriously consider integrating these data to databases of Automatic Landmark Detection Systems (Sadeghian & Kantardzic, 2008).

Background and Relevance

Urban areas now host more than half of the world's population. The United Nations plan a steady increase of urbanization with an urban population - world population ratio expected to reach 60% by 2030. In order to reduce the environmental impact - in particular emissions of greenhouse gas emissions - urban management (waste, energy, water, transport, etc.) must be optimized. The concept of "Smart City" (Hall, 2000) now appears to be the suitable alternative to the "traditional city". Although no clear definition has been proposed to date (Chourabi *et al*, 2012), everyone agrees that the performance improvement of our cities depends largely on the management and the effective exploitation of Information and Communication Technologies (ICT). A smart city is a system based *de facto* on a digital infrastructure and needs consequently relevant data to operate efficiently.

Patterns identified by Noulas *et al* (2011) led them to conclude that geolocated data produced from a Location-Based Social Network (LBSN) such as Foursquare reflected overall daily user activity. Therefore, check-ins appear to be reliable for analysing urban dynamics at the scale of a city, a region or even a country. The work of Bawa-Cavia (2012) on the hot spots, Cheng *et al* (2011) on mobility and Cranshaw *et al* (2012) on the proto-neighbourhood structures provide explicit illustrations. However, social surveys of Cramer *et al* (2011), Evans (2011) and Lindqvist *et al* (2011) showed that there were two main intentions underlying the publication of a check-in: for the user, the main goal is to obtain and keep the titles of mayors and badges he has earned (playfulness aspect). On the other hand, this one wishes to enhance his social status by publishing check-ins only in places he considers strategic. Behind these two motivations, there is a single dimension: the communication or more specifically the geocommunication (Roche *et al*, 2012). In addition, Rost *et al* (2013) indicate that the scope of the check-ins communicative dimension remains, to this day, still undervalued and the postulate that a check-in is equivalent to a visit can be called into question. Insofar as the scope of this communicative dimension is quite difficult to evaluate, we are asking the following question: *Are Foursquare's Check-ins relevant data for the improvement of urban intelligence?*

Furthermore, according to the work of Lynch (1970), landmarks are one of the main reading keys of the city - including paths, edges, districts and nodes - from which we navigate trough space and find our way (i.e. wayfinding). According to the theory of Siegel and white (1975), also known as the acronym LRS, landmarks are, with the Route and the Survey knowledge, the main organization levels of our cognitive map. Golledge (1978) proposed a theory, which is in line with the LRS classification of Siegel and White. This theory of "anchor points" suggests that landmarks of our cognitive maps are hierarchically structured, that is to say that the significance of a reference point differs from one individual to another depending on their lived. Sorrows and Hirtle (1999) proposed much later three categories of landmarks: (1) visual landmarks: these are easily identifiable by their size which contrast with the surrounding buildings. (2) cognitive landmarks: places that fall into this category are generally meaningful to the observer or the community (e.g. a historic site). And (3) the structural landmarks: it is actually highly accessible locations and therefore strategic places because of their position (e.g. buildings located at the intersection of two major traffic axes).

In other words, we argue that improving the mobility of citizens by giving them a landmark-based assistance in real-time contribute to the improvement of urban intelligence; since the smart mobility is one of the main development axes of Smart Cities. Moreover, Schwartz and Naaman (2013) have recently reported that the data stream generated by the social web platforms (i.e. Facebook, Twitter, Foursquare, Flickr, YouTube, Instagram, etc.) could reveal the personal and collective mental maps of cities. Our Ph.D. research focuses on landmarks and Foursquare data. It is therefore an innovative approach because the exploitation of check-ins, from the perspective of spatial cognition - especially wayfinding - has not yet been processed. At most, some researchers have raised the premise that places stored in Foursquare's database were significant points of reference for orientation (cf. Gazzard, 2011; Naaman, 2011; Wakamiya *et al*, 2011; Bentley *et al*, 2012, and Lee *et al*, 2012). Thus, we formulate the following specific research question: *Are geosocial data produced by Foursquare's users a reliable source of information for the identification of landmarks in urban areas?*

Methods and Data

The main goal of this work is to evaluate the representativeness of Foursquare data in terms of landmarks. Because our research thematic has not been explored at all, we will adopt an exploratory approach based on methods that will be both qualitative and quantitative. There will be three main steps. First of all, we plan to collect Foursquare's check-ins at the scale of Quebec City. We will not use the Foursquare's Endpoints API, nor the Twitter's Search or Streaming APIs: we will rather deal directly with GNIP, the authorized Foursquare's data reseller, from whom we will harvest the data over a period of three consecutive months. Once the data are collected, we will proceed with spatial analyses on a GIS in order to identify Quebec City's "main key places". Research works such as those of Lycnh (1970), Allen *et al* (1979), Passini (1984) Couclelis *et al* (1987) and Golledge (1999) showed that there was a close link between landmarks and the concept of neighbourhood. Hirtle (2003) mentions the possibility of using geometric tools such as Voronoi diagrams and Delaunay triangulation (cf. Boots, 1986 Boots et al, 1999). In addition, taking into account the height of buildings as a key indicator, Winter *et al* (2008) have developed a method to generate a hierarchical structure of landmarks based on Voronoi diagrams. We will use their work to generate similar structures based on the geosocial activity of Foursquare's users. Then, we should be able to determine from these areas (1) spatial prototypes (cf. Rosch, 1975 and Huttenlocher *et al*, 1991) of Quebec City, (2) non-prototypes (i.e. places that generate the lowest geosocial activity in the area) and finally (3) intermediate places. In this way, we should obtain a representative sample of places stored in the Foursquare's Database of Quebec City. Finally, we plan to make interviews of those main key places mayors. Most of the survey will be clearly focused on a quantitative approach. Indeed, we know that distances tend to be underestimated from relevant landmarks (e.g. Sadalla *et al*, 1980) and direction estimates appear to be more accurate (Golledge, 1999). Therefore, we will use specific techniques listed by Cauvin (1984), Montello (1991) and Kitchin & Blades (2002) to obtain the estimated distances and directions from Foursquare's mayors; such as the direct magnitude estimation for the estimated distances and the projective convergence (cf. Siegel, 1981 and Kirsac *et* al, 1984) for the estimated directions.

Results

If this research provides evidences of the check-ins reliability, then we could seriously consider integrating these specific kind of data to databases of Automatic Landmarks Detection Systems (ALDSs; Sadeghian & Kantardzic, 2008). The ALDS are tools designed to improve routes by offering to the traveller additional tracking entirely based on landmarks. In summary, ALDSs work as follows: (1) when a route is established, a neighbourhood analysis is performed at each intersection, (2) among all buildings present in the analysed area, the most significant place in terms of landmarks is selected, and (3) the set of landmarks is returned to the traveller as an additional information to consider when approaching an intersection. The criteria for significance of ALDSs are none other than the size of the building, its height and its colour. However, as both authors note, it would be interesting to provide the ALDSs databases attributes such as visiting figures and the historical and cultural significance of each place.

Conclusions

We suggest through this research that if the data produced by geosocial users can be exploited to improve the collective spatial knowledge, then these will contribute to the improvement of urban intelligence. We focus on Location-based social networks (LBSNs) because a check-in is a publication systematically attached to a place. Indeed, place is the smallest spatial unit of the society (Lussault, 2003) on the basis of which we "generate" each of our cognitive reference points (Lynch, 1970). However, these are the main component of our spatial knowledge. It is clearly not a coincidence if the latest version of Google Maps offers a personalized map (Joliveau, 2013) that highlights places of interest for each user according to the activity that he generates on the services of the firm. In addition, the first LBSNs have sprung up there over six years ago, but we did not find any academic research that explored the potential of geosocial data from the perspective of cognitive landmarks. Actually, this potential has been addressed very recently by Schwartz and Naaman (2013). Therefore, we argue that the outcomes of this research will contribute to the advancement of knowledge in the field of geographic information science.

References

Bawa-Cavia, A. (2011). *Sensing The Urban : Using location-based social network data in urban analysis*. Paper presented at the Workshop on Pervasive Urban Applications (PURBA), Ninth International Conference on Pervasive Computing, San Francisco, USA.

Bentley, F., Cramer, H., Basapur, S., & Hamilton, W. (2012). *Drawing the City: Differing Perceptions of the Urban Environment* Paper presented at the CHI'12, Austin, Texas.

Boots, B. N. (1986). *Voronoi (Thiessen) Polygons*. Norwich: Geo Books.

Boots, B. N., Okabe, A., & Sugihara, K. (1999). Spatial tessellations. *Geographical information systems, Vol. 1*, pp. 503-526.

Cauvin, C. (1984). *La perception des distances en milieu intra-urbain : une première approche*. Paris: Éditions du CDSH.

Cheng, Z., Caverlee, J., Lee, K., & Sui, D. S. (2011). *Exploring Millions of Footprints in Location Sharing Services*. Paper presented at the ICWSM '11.

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., Scholl, H. J. (2012). *Understanding Smart Cities: An Integrative Framework*. Paper presented at the Proceedings of 45th Hawaii International Conference on System Sciences.

Couclelis, H., Golledge, R. G., Gale, N., & Tobler, W. (1987). Exploring the anchorpoint hypothesis of spatial cognition. *Journal of Environmental Psychology, Vol. 7*, pp. 99-122.

Cramer, H., Rost, M., & Holmquist, L. E. (2011). *Performing a Check-in: Emerging Practices, Norms and 'Conflicts' in Location-Sharing Using Foursquare*. Paper presented at the MobileHCI 2011, Stockholm, Sweden.

Cranshaw, J., Schwartz, R., Hong, J. I., & Sadeh, N. (2012). *The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City*. Paper presented at the The Sixth International AAAI Conference on Weblogs an Social Media (ICWSM-12), Dublin, Ireland.

Davies, C., Li, C., & Albrecht, J. (2010). Human understanding of space. In M. M. Haklay (Ed.), *Interacting with Geospatial Technologies*: John Wiley & Sons Ltd.

Evans, L. (2011). Location-based services: transformation of the experience of space. *Journal of Location Based Services, Vol. 5*(3 & 4), 20.

Gazzard, A. (2011). Location, location, location: Collecting space and place in mobile media. *Convergence: The International Journal of Research into New Media Technologies, Vol. 17*(No. 4), pp. 405-417.

Golledge, R. G. (1978). Representing, interpreting and using cognized environments. *Papers and Proceedings of the Regional Science Association, Vol. 41*, pp. 169-204.

Golledge, R. G. (1999). *Wayfinding behavior: Cognitive mapping and other spatial processes*. Baltimore: Johns Hopkins Press.

Hall, R. E. (2000). *The vision of a smart city*. Paper presented at the Proceedings of the 2nd International Life Extension Technology Workshop, Paris, France.

Hirtle, S. C. (2003). Neighborhoods and Landmarks. In M. Duckham, Goodchild, M., Worboys, M. (Ed.), *Foundations of Geographic Information Science* (pp. pp. 191-203). London and New York: Taylor & Francis.

Huttenlocher, J., Hedges, H. J., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review, Vol. 98*, pp. 352-276.

Kirasic, K. C., Allen, G. L., & Siegel, A. W. (1984). Expression of configurational knowledge of large-scale environments: student's performance of cognition tasks. *Environment and Behaviour, Vol. 16*(No. 6), pp. 687-712.

Kitchin, R., & Blades, M. (2002). *The cognition of geographic space*. London ; New York: I. B. Tauris.

Lee, R., Wakamiya, S., & Sumiya, K. (2012). Urban area characterization based on crowd behavioral lifelogs over Twitter. *Personal and Ubiquitous Computing, Vol. 17*(No. 4), pp. 605-620.

Lindqvist, J., Cranshaw, J., Wiese, J., Hong, J.I, & Zimmerman, J. (2011). *I'm the Mayor of My House: Examining Why People Use foursquare - a Social-Driven Location Sharing Application*. Paper presented at the Proceedings of the Conference on Human Factors in Computing Systems (CHI-11), Vancouver, BC, Canada.

Joliveau, T. (2013). Chacun sa carte? Le nouveau Google Maps. *Mappemonde, 110*.

Lussault, M. (2003). Lieu. In Belin (Ed.), *Dictionnaire de la géographie et de l'espace des sociétés*. Paris.

Lynch, K. (1970). *The image of the city*. Cambridge: MIT Press.

McNamara, T. P., & Valiquette, C. M. (2004). Remembering Where Things Are. In G. L. Allen (Ed.), *Human Spatial Memory*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Montello, D. R. (1991). The measurement of cognitive distance: Methods and construct validity. *Journal of Environmental Psychology, Vol. 11*, pp. 101-122.

Naaman, M. (2011). Geographic information from georeferenced social media data. *SIGSPATIAL Special, Vol. 3*(No. 2), pp. 54-61.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). *An Empirical Study of Geographic User Activity Patterns in Foursquare*. Paper presented at the Fifth International AAAI Conference on Weblogs and Social Media.

Passini, R. (1984). *Wayfinding in architecture*. New-York: Van Nostrand Rheinhold.

Roche, S., Nabian, K., & Ratti, C. (2012). Are Smart Cities Smart Enough ? In G. A. Press (Ed.), *Spatially Enabling Government, Industry and Citizens : Research development and perspectives* (pp. pp. 215-236).

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, Vol. 7*, pp. 532-547.

Rost, M., Barkhuus, L., Cramer, H., & Brown, B. (2013). *Representation and Communication: Challenges in Interpreting Large Social Media Datasets*. Paper presented at the CSCW ’13, San Antonio, Texas, USA.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology, Vol. 6*(No. 5), pp. 517-528.

Sadeghian, P., & Kantardzic, M. (2008). The New Generation of Automatic Landmark Detection Systems: Challenges and Guidelines. *Spatial Cognition & Computation: An interdisciplinary journal, Vol. 8*, pp. 252-287.

Schwartz, R., & Naaman, M. (2013). *Collective and Individual Mental Maps of the City in Social Awareness Streams*. Paper presented at the GeoHCI Workshop at CHI 2013, Paris, France.

Siegel, A. W. (1981). The externalization of cognitive maps by children and adults: in search of ways to ask better questions. In L. S. Liben, A. H. Pattterson & N. Newcombe (Eds.), *Spatial Representation and Behavior Across the Life Span*. New York: Academic.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large scale environments. In W. H. Reese (Ed.), *Advances in child development and behavior* (Vol. Vol. 10). New York.

Sorrows, M. E., & Hirtle, S. C. (1999). *The Nature of Landmarks for Real and Electronic Spaces*. Paper presented at the Proceedings of the International Conference on Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science.

Stefanidis, A., Crooks, A., & Radzikowski, J. (2011). Harvesting ambient geospatial information from social media feeds. *GeoJournal*, 20 p.

Wakamiya, S., Lee, K., & Sumiya, K. (2011). *Crowd-based Urban Characterization: Extracting Crowd Behavioral Patterns in Urban Areas from Twitter*. Paper presented at the ACM LBSN '11, Chicago, USA.