

On the Way to Platial Analysis: Can Geosocial Media Provide the Necessary Impetus?

Proceedings of the
First Workshop on Platial Analysis



20–21 September
Heidelberg, Germany

René Westerholt – Franz-Benjamin Mocnik – Alexander Zipf
(Editors)

Please Cite This Volume As

R Westerholt, F-B Mocnik, and A Zipf (eds., 2018): *On the Way to Platial Analysis: Can Geosocial Media Provide the Necessary Impetus? Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*

<https://doi.org/10.5281/zenodo.1475269>

Location and Dates

Heidelberg, Germany; 20–21 September 2018

Convenors

René Westerholt (Heidelberg University, Germany)

Franz-Benjamin Mocnik (Heidelberg University, Germany)

Alexander Zipf (Heidelberg University, Germany)

Keynote Speakers

Clare Davies (University of Winchester, United Kingdom)

Alexis Comber (University of Leeds, United Kingdom)

Programme Committee

Gennady Andrienko (City University London, United Kingdom)

Thomas Blaschke (University of Salzburg, Austria)

Dirk Burghardt (Technical University of Dresden, Germany)

Alexis Comber (University of Leeds, United Kingdom)

Sara Irina Fabrikant (University of Zurich, Switzerland)

Andrew U Frank (TU Wien, Austria)

Hans Gebhardt (Heidelberg University, Germany)

Michael F Goodchild (University of California, Santa Barbara, United States)

Krzysztof Janowicz (University of California, Santa Barbara, United States)

Alan MacEachren (The Pennsylvania State University, United States)

Grant McKenzie (McGill University, Canada)

Franz-Benjamin Mocnik (Heidelberg University, Germany)

Alenka Poplin (Iowa State University, United States)

João Porto de Albuquerque (University of Warwick, United Kingdom)

Ross Purves (University of Zurich, Switzerland)

Simon Scheider (Utrecht University, The Netherlands)

Lisa Teichmann (McGill University, Canada)

Sabine Timpf (University of Augsburg, Germany)

René Westerholt (Heidelberg University, Germany)

Stephan Winter (University of Melbourne, Australia)

Diedrich Wolter (University of Bamberg, Germany)

Alexander Zipf (Heidelberg University, Germany)

EDITORIAL

pp. 1–5

Introduction to the PLATIAL'18 Workshop on Platial Analysis

R Westerholt, F-B Mocnik, and A Zipf

INVITED PAPERS

pp. 7–14

Quantitative Platial Analysis: Methods for Handling and Representing
Platial Heterogeneity and Linking Varying Concepts of Place

A Comber, A Butler, N Malleson, and A Schafran

pp. 15–20

Place and Placing Locations: A Cognitive Perspective

C Davies

CONCEPTUAL ANATOMY OF PLACE

pp. 21–27

Pinpointing Dream Settings onto Place Cookies

CM Iosifescu Enescu and L Hurni

pp. 29–36

The Near-Decomposability Paradigm Re-Interpreted for Place-Based GIS

T Blaschke and ST Piralilou

DISCLOSING PLACES FROM HUMAN DISCOURSE

pp. 37–43

**Turin's Foodscapes: Exploring Places of Food Consumption
Through the Prism of Social Practice Theory**

A Calafiore, G Boella, E Grassi, and C Schifanella

pp. 45–52

Digital Imaginations of National Parks in Different Social Media: A Data Exploration

V Heikinheimo, H Tenkanen, T Hiippala, and T Toivonen

BRIDGING SPACE AND PLACE

pp. 53–59

From Space to Place and Back Again: Towards an Interface Between Space and Place

E Papadakis, G Baryannis, and T Blaschke

pp. 61–65

The Value of Detours

SN Vardag and S Lautenbach

EXPLORATORY AND VISUAL ANALYTICS OF PLACE

pp. 67–73

A Contribution to the Visualization of the Diversity of Places

M Gröbe and D Burghardt

pp. 75–82

Data Mining of Network Events With Space-Time Cube Application

V Putrenko, N Pashynska, and S Nazarenko

Introduction to the PLATIAL'18 Workshop on Platial Analysis

– Editorial –

Rene Westerholt^{1,2}, Franz-Benjamin Mocnik², and Alexander Zipf²

¹*Centre for Interdisciplinary Methodologies, University of Warwick, UK*

²*Institute of Geography, Heidelberg University, Germany*

The concept of “place” is about to become one of the major research themes in the discipline of geographical information science (GIScience), as well as in adjoining fields. Briefly put, while locations provide objective references (e. g., point coordinates), places are the units utilized by humans to approach the geographic world (Goodchild, 2015). On the one hand, the current “platial turn” in GIScience is caused by the plethora of oftentimes user-generated and particularly urban geographic datasets, which have become available in the last years (e. g., geosocial media feeds). These so-called ambient geospatial datasets (Stefanidis et al., 2013) mirror digitally small and limited glimpses of the everyday lives of people, and how they approach and experience the geographic world. Ambient geographic datasets may thus be understood as something deeper than just mere “attributes referenced over point locations”, which is why they have recently been conjectured to be of platial rather than of spatial nature (Quesnot and Roche, 2015). “Platial” can hereby be understood as the place-based counterpart to the space-based adjective “spatial”.

Understanding these either individual or collective, digitally collected experiences requires taking account of social, cultural, behavioural, and cognitive aspects. This endeavour therefore opens up a delicate opportunity for interdisciplinary collaboration, transcending disciplinary boundaries. Alongside this, researchers have become recently aware of the limitations of a purely spatial notion of GIS. Despite its undoubted success over the last decades, what spatial GIS effectively does when investigating human data is facilitating complex affairs into rather simplistic geometric primitives like points, lines, and polygons. These units might be convenient to work with, but they are not fully sufficient for addressing deeply human-geographic and social-scientific questions. For these reasons, researchers have recently called for a paradigm change towards a platial counterpart to the established spatial notion of GIS and quantitative analysis (Goodchild, 2011, 2015; Stedman, 2003), allowing to represent and analyse platial information by computing machinery. This will ultimately allow geographical, sociological, and other related scholars to support their studies by more realistic quantitative inferences.

The PLATIAL'18 workshop makes a significant contribution towards these developments and is meant to be the starting point for a series of future events. What sets this workshop apart from others dealing with the concept of place is that the focus is decisively on its quantitative investigation and conceptual formalization. Nevertheless, PLATIAL'18 accommodates a wide range of aspects all of which in one or another way are related to the two outlined core foci. This is well reflected by the various topical sessions into which the workshop has been organized. These include “Conceptual Anatomy of Place”, “Disclosing Places from Human Discourse”, “Bridging Space and Place”, and “Exploratory and Visual Analytics of Place”. This topical variety, on the one hand, reflects the breadth of the concept of place, but, on the other hand, also the early stage at which we still are from a GIScience point of view. The sessions also demonstrate the success of the workshop in bringing together scholars from a range of different disciplines to work together towards a platial notion of analysis. The following

R Westerholt, F-B Mocnik, and A Zipf (2018): *Introduction to the PLATIAL'18 Workshop on Platial Analysis*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 1–5

<https://doi.org/10.5281/zenodo.1475267>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

paragraphs are indicative for this success. They summarize the above mentioned sessions and give brief summaries of the contributions accepted for oral presentation.

The content sessions were concomitantly inspired by two keynote talks. One talk was given by Alexis Comber (University of Leeds), who reported about platial heterogeneity and linking various place concepts (Comber et al., 2018). A second talk emphasizing a cognitive perspective of place (Davies, 2018) was delivered by Clare Davies (University of Winchester). Both of these talks touch upon very important and fundamental aspects of place-based analysis. Alexis highlighted the importance of the distinction we have to make between “space” and “place” when it comes to quantitative analysis. In his talk, he utilized the example of denigrated places (people assigning the term “shithole”; Butler et al. 2018), for which almost no easily interpretable spatial pattern is found. The results presented, however, demonstrate that insightful patterns can be found when taking a platial perspective instead. This shows that the spatial framework is limited when it comes to subjective platial information, confirming empirically experimental results from the literature (e. g., Westerholt et al. 2016). Denigrating places is also largely related to human cognition. Indeed, cognition is of particular importance to user-generated datasets like geosocial media or the mapping project OpenStreetMap, the latter of which is based on a folksonomy heavily influenced by mental conceptualizations of people (Mocnik et al., 2017). In her talk, Clare emphasized the importance of human cognition as an integral part of GIScience (Montello and Mark, 2018), which is particularly important to the study of places. She reported about the role places have for categorizing related locations, and how the concept of place might thus be understood as a classification heuristic. Both keynote talks have been highly inspirational and stimulated discussions throughout the workshop.

The session that deals with the core of the concept of place is entitled “Conceptual Anatomy of Place”. The contribution made by Blaschke and Piralilou (2018) forms part of this and deals with the inherent complexity of place. According to Simon (1977, Chapter 4.4), all viable systems are (near-)decomposable into their constituent parts, no matter whether they are of social, technical, or physical nature. For this reason, Blaschke and Piralilou (2018) hypothesize that this might also be the case with places. In order to cope with complexity, and also with scaling issues, they further propose transferring concepts from object-based image analysis to the analysis of places. A second paper allocated to this session explores ways to formalize the relations between places experienced in dreams, with those experienced consciously while awake. Iosifescu Enescu and Hurni (2018) propose the concept of a layered so-called “place cookie” for this purpose, which can be used to classify places with respect to their familiarity to a dreamer. The place cookie concept also allows to combine spatial with platial notions distance through linking the cookie back to geographical space. Overall, this session tackles two different but related topics, covering very fundamental aspects of the nature of place and their investigation.

Verbalization is a key factor to the investigation of places (Goodchild, 2011; Winter and Freksa, 2012). The way we have access to places is mostly through considering verbalized expressions made by people. For this reason, our second session is dedicated to the extraction of place-based information from human discourse. One approach to this is presented by Calafiore et al. (2018), who work on the case of food consumption in Turin, Italy. They extract shared notions of place related to how people experience the “foodscape” of the city by investigating crowdsourced TripAdvisor data. Using clustering techniques and by applying social practice theory, the case study reveals links between socially-defined groups and jointly experienced places. In a related yet slightly different manner, Heikinheimo et al. (2018) investigate how well different geosocial media feeds are actually suited to be used for disclosing place-based digital imaginations. Adopting a Finnish national park use case, the authors compare information from Flickr, Instagram, and Twitter. Thereby, they review these with respect to their information content, originality of locational information, and further factors. This session largely reflects the empirical exploration of places, which is a very important cornerstone on the way towards evidence-based platial research.

Place has frequently been described as space infused with meaning (Tuan, 1977). Based on this notion, our fourth session aims to link the two universes of “space” and “place”. Assuming an inherent link between space and place, Papadakis et al. (2018) present a philosophical contribution towards bridging these two paradigms. They present first approaches to an interface that, in a reciprocal manner, allows to convert between space and place by utilizing different kinds of intermediary spaces as introduced by Couclelis (1992). Another approach is followed by Vardag and Lautenbach (2018), who investigate the relationships between the geometric length of detours (spatial) and the associated

additional personal values experienced through considering these instead of shortest paths (platial). For this purpose, they utilize (semi-)automated methods to extract semantic links from georeferenced assessments of peoples' moods collected in an in situ manner.


The fourth session of the workshop is devoted to exploratory and visual analytics of place. The contribution made by Gröbe and Burghardt (2018) proposes the cartographic technique of micro-diagrams to be used for visualizing the diversity attached to places. In essence, this technique entails the generation of mapped diagrams enabling to represent the thematic (or any related kind of) diversity of places. In contrast to this cartographic approach, Putrenko et al. (2018) make use of space-time cubes, established by time geography, to explore the relations between social phenomena and locations. This way, and by additionally applying spatial-statistical measures, it is possible to indicate place-related events from social networks.


The workshop conducted this year portrays an impressive breadth, reflecting the diversity that is inherent to the concept of place. It also, however, unveils the lack of clarity in how geospatial and related scholars refer to and deal with the notion of place. The PLATIAL'18 workshop contributes to the consolidation of this latent and widespread vagueness. In this vein, the workshop is in line with other events carried out in 2018, for instance, a session dedicated to “place” organized at the GIScience conference held in Melbourne, Australia. It will be interesting to see in which directions the platial turn in GIScience will develop in the upcoming years. We are looking forward to forming part in this exciting endeavour through continuing the PLATIAL series with another fruitful PLATIAL'19 event to be held next year.


Acknowledgements

We are grateful to everyone who contributed to making this workshop a big success! Namely, we want to thank Saskia Rupp and Johanna Schwehn (student assistants), as well as Bettina Knorr and Angelika Hoffer (administrators) for their invaluable help behind the scenes. We also feel very much obliged to our keynote speakers Clare Davies and Alexis Comber for their extremely inspiring talks. Thanks also go to the external panelists who, with their forward-looking statements on the further development of the topic, have decisively stimulated all the participants present to continue working on “Place” beyond the workshop. These include Thomas Blaschke, Dirk Burghardt, Alexis Comber and Clare Davies. Further we have to thank the programme committee for their excellent reviews of our submissions: Gennady Andrienko, Thomas Blaschke, Dirk Burghardt, Alexis Comber, Sara Irina Fabrikant, Andrew U Frank, Hans Gebhardt, Michael F Goodchild, Krzysztof Janowicz, Alan MacEachren, Grant McKenzie, Alenka Poplin, João Porto de Albuquerque, Ross Purves, Simon Scheider, Lisa Teichmann, Sabine Timpf, Stephan Winter, and Diedrich Wolter. Last but not least, we are grateful to all participants of the workshop for making PLATIAL'18 a tremendous success!

ORCID

Rene Westerholt  <https://orcid.org/0000-0001-8228-3814>

Franz-Benjamin Mocnik  <https://orcid.org/0000-0002-1759-6336>

Alexander Zipf  <https://orcid.org/0000-0003-4916-9838>

References

Blaschke, Thomas and Piralilou, Sepideh Tavakkoli: *The near-decomposability paradigm re-interpreted for place-based GIS*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 29–36. doi: 10.5281/zenodo.1472741

Butler, Alice; Schafran, Alex; and Carpenter, Georgina: *What does it mean when people call a place a shithole? Understanding a discourse of denigration in the United Kingdom and the Republic of Ireland*. *Transactions of the Institute of British Geographers*, 43(3), 2018, 496–510. doi: 10.1111/tran.12247

Calafiore, Alessia; Boella, Guido; Grassi, Elena; and Schifanella, Claudio: *Turin's foodscapes: exploring places of food consumption through the prism of social practice theory*. In: Westerholt, Rene; Mocnik,

Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 37–43. doi: 10.5281/zenodo.1472743

Comber, Alexis; Butler, Alice; Malleson, Nick; and Schafran, Alex: *Quantitative platial analysis: methods for handling and representing platial heterogeneity and linking varying concepts of places*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 7–14. doi: 10.5281/zenodo.1472735

Couclelis, Helen: *People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS*. Proceedings of the International Conference GIS, From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning, 1992, 65–77. doi: 10.1007/3-540-55966-3_3

Davies, Clare: *Place and placing locations: a cognitive perspective*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 15–20. doi: 10.5281/zenodo.1472737

Goodchild, Michael F: *Formalizing place in geographic information systems*. In: Burton, Linda M; Matthews, Stephen A; Leung, ManChui; Kemp, Susan P; and Takeuchi, David T (eds.), *Communities, neighborhoods, and health*, New York, NY: Springer, 2011. 21–33. doi: 10.1007/978-1-4419-7482-2_2

— *Space, place and health*. *Annals of GIS*, 21(2), 2015, 97–100. doi: 10.1080/19475683.2015.1007895

Gröbe, Mathias and Burghardt, Dirk: *A contribution to the visualization of the diversity of places*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 67–73. doi: 10.5281/zenodo.1472751

Heikinheimo, Vuokko; Tenkanen, Henriikki; Hiippala, Tuomo; and Toivonen, Tuuli: *Digital imaginations of national parks in different social media: a data exploration*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 45–52. doi: 10.5281/zenodo.1472745

Iosifescu Enescu, Cristina Mihaela and Hurni, Lorenz: *Pinpointing dream settings onto place cookies*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 21–27. doi: 10.5281/zenodo.1472739

Mocnik, F.-B; Zipf, Alexander; and Raifer, Martin: *The OpenStreetMap folksonomy and its evolution*. *Geo-spatial Information Science*, 20(3), 2017, 219–230. doi: 10.1080/10095020.2017.1368193

Montello, Daniel R and Mark, David M: *The world, the computer, and the mind: how Andrew Frank helped make human language and cognition cornerstones of geographic information science*. *International Journal of Geographical Information Science*, 32(12), 2018, 2535–2550. doi: 10.1080/13658816.2018.1436714

Papadakis, Emmanuel; Baryannis, George; and Blaschke, Thomas: *From space to place and back again: towards an interface between space and place*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 53–59. doi: 10.5281/zenodo.1472747

Putrenko, Viktor; Pashynska, Nataliia; and Nazarenko, Sergii: *Data mining of network events with space-time cube application*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 75–82. doi: 10.5281/zenodo.1472753

Quesnot, Teriitutea and Roche, Stéphane: *Platial or locational data? toward the characterization of social location sharing*. Proceedings of the 48th Hawaii International Conference on System Sciences, 2015, 1973–1982. doi: 10.1109/HICSS.2015.236

Simon, Herbert A: *Models of discovery. And other topics in the methods of science*. Dordrecht, The Netherlands: Springer, 1977. doi: 10.1007/978-94-010-9521-1

Stedman, Richard C: *Sense of place and forest science: toward a program of quantitative research*. Forest Science, 49(6), 2003, 822–829

Stefanidis, Anthony; Crooks, Andrew; and Radzikowski, Jacek: *Harvesting ambient geospatial information from social media feeds*. GeoJournal, 78(2), 2013, 319–338. doi: 10.1007/s10708-011-9438-2

Tuan, Yi-Fu: *Space and place. The perspective of experience*. Minneapolis, MN: University of Minnesota Press, 1977

Vardag, Sanam Noreen and Lautenbach, Sven: *The value of detours*. In: Westerholt, Rene; Mocnik, Franz-Benjamin; and Zipf, Alexander (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*. 2018, 61–65. doi: 10.5281/zenodo.1472749

Westerholt, Rene; Steiger, Enrico; Resch, Bernd; and Zipf, Alexander: *Abundant topological outliers in social media data and their effect on spatial analysis*. PLoS ONE, 11(9), 2016, e0162360. doi: 10.1371/journal.pone.0162360

Winter, Stephan and Freksa, Christian: *Approaching the notion of place by contrast*. Journal of Spatial Information Science, 5, 2012, 31–50. doi: 10.5311/JOSIS.2012.5.90

Quantitative Platial Analysis: Methods for Handling and Representing Platial Heterogeneity and Linking Varying Concepts of Place

– Invited Keynote Paper –

Alexis Comber^{ORCID}, Alice Butler^{ORCID}, Nick Malleison^{ORCID}, and Alex Schafran^{ORCID}

School of Geography, University of Leeds, UK

This paper explores potential approaches for quantitative platial analysis. It revisits some of the early work examining place in social media data in light of recent proposals for a platial GIS. Focussing on Massey's concept of space that incorporates a sense of belonging and kinship, where space becomes place through social relations, it uses coded Twitter data containing the term "shithole" to generate a predictive models of different types of platial denigration. These are used to infer the spatial distribution of different types of platial denigration. The results show that there is little spatial pattern to denigration of different places and sports facilities, but that denigration of one's own local area and of one's own personal space have highly localized distributions. The discussion indicates a number of areas for further research with a particular warning against developing platial GISs as has been suggested by many authors. Other explicitly GIScience avenues may be more productive and insightful.

Keywords: spatial analysis; platial analysis; Twitter; shithole

1 Introduction

This introduction briefly covers the concept of place in geography and then the inherent social construction of spatial data in more detail. These lay the ground for a critique of how the GIScience/spatial analysis/geocomputation community have hitherto sought to take a platial turn, and set up the platial analysis later in the paper.

The concept of place is a core consideration in critical geography. It has a number of characteristics that GIScientists have struggled to robustly accommodate within a *platial* information system: place refers to multiple spatial concepts; places are spaces where the notion of distance is irrelevant; and place defines the socio-cultural context in which everyday lives are lived. Doreen Massey, offers a useful conceptualization of the idea of "place", that emphasizes the changing nature of place and place-making (Massey, 2000). She clearly and concisely defines what constitutes a place: "places are spaces of social relations" (Massey, 2000, p. 459). This definition highlights the key difference between space and place and highlights the centrality of social relations for place-making. In this the concept of "place" incorporates a sense of belonging and kinship where space is anonymous but has the potential – with the introduction of social relations – to become place. Places also evolve and are dynamic: "the place goes on being made" (Massey, 2000, p. 464) emphasizing the force of time and the necessity of understanding the entirety of a place rather than at a "snapshot" moment in time. Place is also relative to the multiple and differentiated and public: "one place' can be known in numerous ways" (Massey, 2000, p. 464), suggesting that, in the same way that time changes a place so, too, does one's social

A Comber, A Butler, N Malleison, and A Schafran (2018): *Quantitative Platial Analysis: Methods for Handling and Representing Platial Heterogeneity and Linking Varying Concepts of Place*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18), pp. 7–14

<https://doi.org/10.5281/zenodo.1472735>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

relations with that place. There is a strong nostalgic element to this concept of place, and for Massey impersonal, symbolically distant “space” becomes “place” when social relations and memories are interwoven with physical space. Thus place-making is ongoing and relates to belonging. A place is known and intimate and depends on the imposition of bodies and relations between those bodies in order to exist as anything other than space. Massey’s formulation of space and place offer to this study an understanding that foregrounds the process of place: place is ongoing and continually shaped.

How we represent the real world in our spatial databases is a key concept in geography as it determines the nature of the questions we are able to answer in our geographical and spatial analyses. Consequently, notions of place have also been extensively considered within the broad domain of GIScience and a long-standing corpus of research exists. As GIS started to emerge into the mainstream in the early 1990s, many researchers started to critically examine how we were using the technology and digital data, particularly how real world features were delineated and encoded in spatial databases. The work of Barry Smith, David Mark, and Andrew U. Frank are exemplars in this area. They were concerned with the concepts and meaning that are implicitly embedded in data, how features were delineated, their labels (“lake” vs. “lac”), the semantics, culture and philosophies they represented, and how to appropriately encode them in our databases. There are three key and interlinked but forgotten research ideas from this time that worth revisiting:

1. The contested nature of features through the notion of *fiat* and *bona fide* objects and boundaries (Burrough and Frank, 1996; Smith, 1995, 2001; Smith and Varzi, 2000). In brief, *fiat* objects exist only because of some kind of cognitive demarcation and owe their existence to acts of human decision. *Bona fide* objects do not and are independent of human conceptions. An example *fiat* might be representational choices over where to place the forest boundary as trees intergrade with shrub land cover in successional vegetation. Such choices are routinely made in the construction of all spatial data and inevitably have the potential to result in analytical variation and therefore uncertainty;
2. Acknowledgement that different groups of people conceptualize the world in different ways. The names and labels we give to things, places and geographic phenomena reflect group perceptions of characteristics (Mark and Turk, 2003; Smith and Mark, 2003) due to linguistic and cultural factors (Smith and Mark, 1998). This recently been observed in crowdsourced data (Comber et al., 2016) and data from different groups has been shown to result in significantly different results when used in analysis; and
3. Geographic objects or processes and their group meanings also vary fundamentally with scale (Fisher et al., 2004).

In the early days of GIS/GIScience researchers were fundamentally concerned with these core representational issues and the uncertainties that might occur when, e. g., data and user perceptions of an object differ. These considerations persist and imply that spatial data will always be socially constructed (Harvey, 2000). They also lead to a health warning that has largely been ignored as the geo-digital revolution: geographic entities are inherently and intimately connected to the space that they occupy and to the manner of their (human) conceptualization (Varzi, 2001).

These issues, fundamental to all spatial data and for spatial data analysis, have largely been overlooked by the GIScience community (COSIT excepted) in recent years. This is mostly due to the nature of digital information systems and the ease now with which we are able to collect, process, and analyse spatial data of all kinds. Our situation is analogous to the old joke “What is a lecture?”¹.

It is also reflected in recent forays (like this one) considering how GIScience and other information technologies could take a platial turn. These have been driven by the opportunities for digital place-based research afforded by technological developments. A number of papers have generally made the following points (extracted variously from text) (Gao et al., 2013; Goodchild, 2011, 2015; Roche, 2016):

1. Places are messy and difficult to define and pin down. Places are poorly defined in “space”, frequently with indeterminate boundaries, and the individual perceptions of those places and their properties vary. However, for GIS they need to be “identifiable” to exist and consequently named places, despite frequently being vaguely defined and context dependent, are used to provide the link between (Euclidian, mapped) space and behavioural place, because place-names can be converted to coordinates.

2. GIS is not very good at representing place. Although GI technologies are inherently spatial, they can be used as a mediating object in planning and to support citizen engagement. But they are not very good at handling alternative (i. e., non-Euclidean) or ambiguous representations of places. This is because places within any human discourse may be vaguely defined and context dependent in contrast to the precise and objective coordinates of space.
3. Actually, beyond understanding behaviours through place names we do not know how to do this. New personal digital GPS-enabled technologies provides opportunities for a new relationship between GIScience and place through social media, VGI, geoweb, etc. This would support spatial enablement, literacy, and empowerment but new theories are needed for such geoplatial methods, technics, and tools.

Integrating platial and spatial in this manner may be mis-guided. They can certainly be linked (see the analysis below) but each has their unique strictures and rubrics. Robust theories take time to evolve, the technology is moving faster than the thinking, and we still have basic things that need to be addressed in GIScience (e. g., we can't even deal with time very well). Spatial detective work is at the core of the role spatial analysis/GIScience: our job is to help domain experts understand what is going in their place (and our space) by developing methods for (s)p(l)atial explorations of data and processes.

There alternatives to the “we need a new philosophy-geoplatial turn” route: returning to previous research we can see that many platial-spatial paths that we have been discarded without being exhausted – mainly because of the technological and data neophilia that has pervaded information sciences in recent years. One such direction is the work by GIScientists in the early 2000s that explicitly sought to link to descriptions and mappings of place to construct alternative gazetteers, vernacular geographies, to create maps of “places”, etc. using text and sentiment mining of place names in Flickr tags, alternative POIs, and citizen participatory mapping. New forms of data were becoming available and provided opportunities to understand how people were experiencing their environment. This activity acknowledged the fact that notions of place *are* grounded semantics, meaning, (spatial) cognition, perception, and linguistics, which researchers sought to capture. It also acknowledged the inherent relativity of place and sought to construct multiple geographies. It recognized that any given palatial concept may be understood differently by different people, groups, and cultures. Such approaches provide a framework (rather than a philosophical and ontological tautology) for GIScientists to work with platial data. It is illustrated in the next section through a spatial analysis of where people use the term “shithole” in relation to different “places”.

2 Data and Methods

Butler et al. (2018) examined the use and intended meaning of the term “shithole” in tweets to understand place based stigma and how discourses of denigration are shaped by the availability and uptake of social media platforms. They found that in most cases, “shithole” or “#shithole” was used to refer to places that the tweeter was not from and that some uses reflected a desire to leave (predominantly contributed by female users). The geography of these tweets had no specific pattern, leading the authors to note that “individuals cry for help and want to leave virtually everywhere and every type of place. They want to leave dirty, ill-equipped homes, villages that are boring, towns that lack amenities, and cities that are dirty and full of ‘others’” (Butler et al., 2018, p. 11).

This paper uses Butler’s data to create predictive spatial models. The data included the following relational geography codes to indicate the scale, scope, and target of each tweet:

- **Other** when the tweet referred to a place that was not the home of the tweeter;
- **Own** where the tweet referred to the tweeter’s own area or region;
- **Facilities** such as a sports stadium;
- **Personal** which usually was used to refer to a room, home, or place of work.

The 1989 tweets containing “shithole” were stemmed, creating a corpus of 2653 stemmed terms and used to train an elastic net/lasso (ELN) model. The model was then applied to a larger, uncoded,

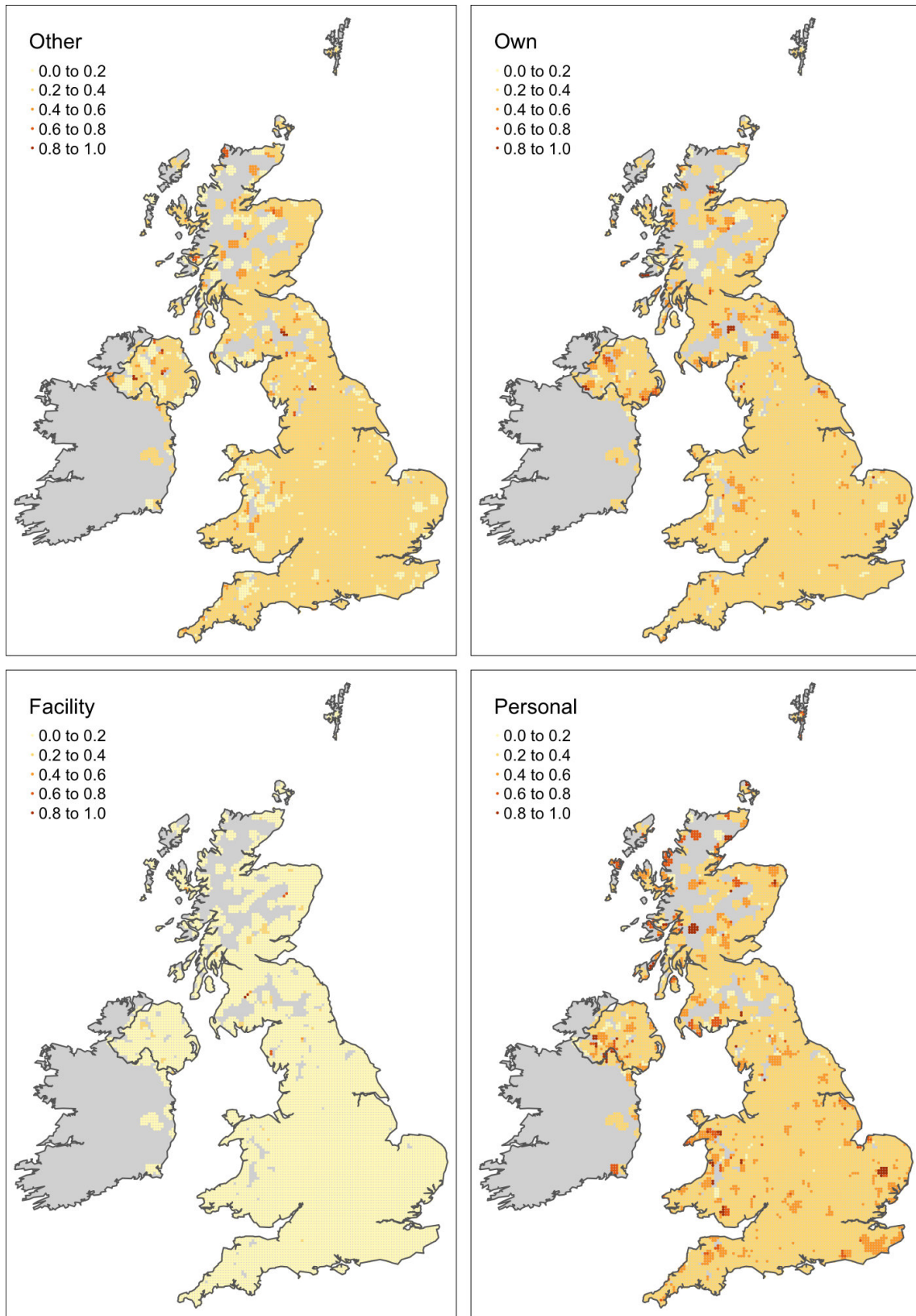


Figure 1: The spatial distribution of tweets classified into the different relation geography classes of “shithole”: the mean posterior probability of each class.

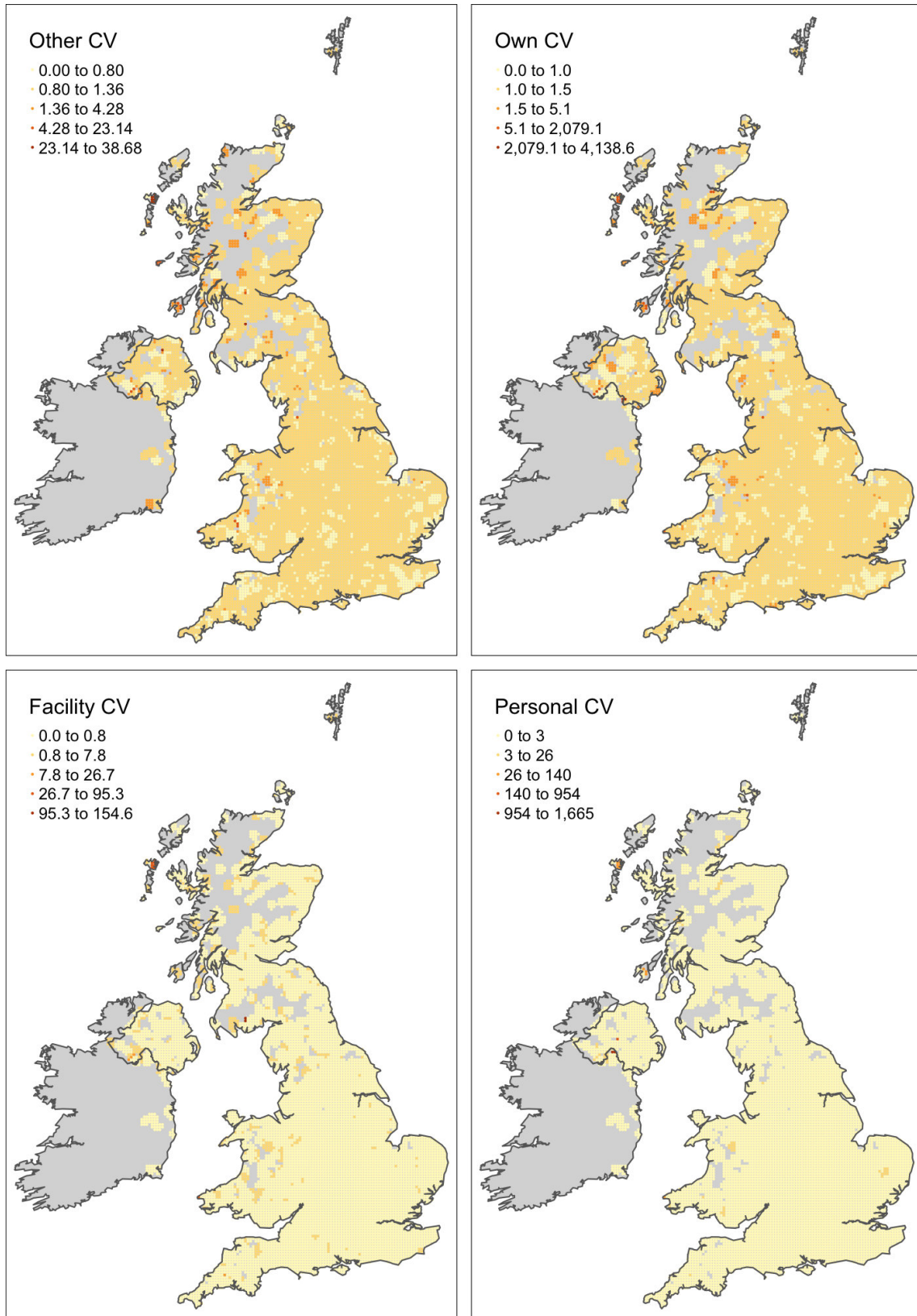


Figure 2: The spatial variation of the mean posterior probability of each class.

dataset of approximately 1 million geo-located (not geo-tagged) tweets and used to predict the relational geography class. The spatial distribution of each class forms the basis of the analysis reported in this paper. The aim, however, was not just to generate a hard, crisp classification of tweets, but to explore the underlying geographies of variations in the strength of classification through the model posterior probabilities and their variances. These soft classification measures provide a route to understand spatial variation in the way that the term “shithole” is used in different parts of the country.

3 Results

The ELN models was constructed from a binary sparse matrix of stemmed terms, indicating the presence of a given term. In-sample model fits showed the model to correctly predict the class of “shithole” 94.3% of the time. The model was then used to predict the 4 classes for each of the 961,597 tweets. These were summarized and smoothed over a 5 km grid using a 10 km radius circular window. For each location, the mean posterior probabilities for each class from the ELN model were calculated along with the coefficient of variation. The first gives an indication of the general trend and the second indicates the heterogeneity of that value. The maps of these are shown in Figures 1 and 2 for each class.

The maps in Figure 1 show a number of things, some of which were identified by text (Butler et al., 2018):

- **Other** has no discernible geographic pattern indicating that tweeters everywhere are equally likely to denigrate a different place.
- **Own** has strong local concentrations in the Scottish borders, Northern Ireland and Mid-Wales indicating higher levels of expressions of denigration by residents in these areas of their own locality than others.
- **Facility** has an even distribution as Butler stated – simply, there is no geography.
- **Personal** is heavily found at the geographic fringes in rural remote places. This suggests that in these places, tweets tend to be closer to denigration of personal space (room, home, or place of work) than any other of the classes. People do not like their remote lives and social media gives them an opportunity to express that.

The maps in Figure 2 indicate that the variation in pooled posterior probability is relatively even for each of the four classes, with a few notable exceptions:

- **Other** has some notable pockets of high variation in the Hebrides and Grampians in Scotland, Cork in Ireland, and Mid Wales.
- **Own** has similar localized trends to **Other**.
- There are no obvious trends in the variation of **Facility** or **Personal**.

Places with low variation in Figure 2 indicate that the mean value in Figure 1 is highly representative of the tweets in that area.

4 Discussion

This research used a standard technique to model spatial variation in perceptions of place, and by developing a predictive model from data that had been manually coded for platial characteristics. This was applied to other Twitter data, for which the geography was known. The geography was derived from the Twitter users’ home location and not the target of their tweets. So, e.g., areas with high posterior probabilities for “Other” are not locations that are denigrated by people who do not live there. Rather it indicates a greater probability for people in that region being more likely to denigrate other places in their tweets.

The analysis used a 5 km grid to summarize the typical values of the classified tweets and their variations. Summarizing and visualizing data in this way provides a useful starting point for discussions

with domain experts – there may be well known social gradients that are described by these mapped distributions. The mean posterior probabilities provide an indication of the multiple potential places that may be present at any location. Soft representations such as these (and, e. g., fuzzy sets) allow alternative and multiple representations of the same geographic phenomena, ones that allow different perceptions of space and thereby place to be accommodated (e. g., Comber and Kuhn 2018).


Finally, there are lots of areas for further work that will be expanded in the full journal paper arising from the publication: alternative classification models (initial work showed quite different results for “Personal” using linear discriminant analysis rather ELN), the use of medians and inter-quartile ranges rather than means and coefficients of variation to quantify central tendencies in a way that is resistant to numerical outliers, deeper investigation of specific locales to try to unpick and understand the local trends that have been observed here, and testing the sensitivity of the processes captured in the predictive models to scale and how the patterns and trends observed vary over different units of analysis. Finally it would be useful to evaluate the representativeness of the platial phenomena captured by the classified twitter data from some alternative data source.

This kind of research avenue is likely to be more productive and will better support platial analysis than ones that seek to steer GIScience towards a platial turn: we should stick to what we are good at – representation, scale, and uncertainty – areas that no other disciplines can do as well as us.


Notes


1. Answer: It is the process by which the lecturer’s words are transferred to the student’s notes without going through the brain of either.

ORCID

Alexis Comber  <https://orcid.org/0000-0002-3652-7846>

Alice Butler  <https://orcid.org/0000-0002-7205-9832>

Nick Malleson  <https://orcid.org/0000-0002-6977-0615>

Alex Schafran  <https://orcid.org/0000-0003-1990-925X>

References

Burrough, Peter A and Frank, Andrew U (eds.): *Geographic objects with indeterminate boundaries*. London, UK: CRC, 1996

Butler, Alice; Schafran, Alex; and Carpenter, Georgina: *What does it mean when people call a place a shithole? Understanding a discourse of denigration in the United Kingdom and the Republic of Ireland*. *Transactions of the Institute of British Geographers*, 43(3), 2018, 496–510. doi: 10.1111/tran.12247

Comber, Alexis and Kuhn, Werner: *Fuzzy difference and data primitives: a transparent approach for supporting different definitions of forest in the context of REDD+*. *Geographica Helvetica*, 73(2), 2018, 151–163. doi: 10.5194/gh-73-151-2018

Comber, Alexis; Mooney, Peter; Purves, Ross S; Rocchini, Duccio; and Walz, Ariane: *Crowdsourcing: it matters who the crowd are. the impacts of between group variations in recording land cover*. *PloS ONE*, 11(7), 2016, e0158329. doi: 10.1371/journal.pone.0158329

Fisher, Peter; Wood, Jo; and Cheng, Tao: *Where is Helvellyn? fuzziness of multi-scale landscape morphometry*. *Transactions of the Institute of British Geographers*, 29(1), 2004, 106–128. doi: 10.1111/j.0020-2754.2004.00117.x

Gao, Song; Janowicz, Krzysztof; McKenzie, Grant; and Li, Linna: *Towards platial joins and buffers in place-based GIS*. *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Computational Models of Place (COMP)*, 2013, 42–49. doi: 10.1145/2534848.2534856

Goodchild, Michael F: *Formalizing place in geographic information systems*. In: Burton, Linda M; Matthews, Stephen A; Leung, ManChui; Kemp, Susan P; and Takeuchi, David T (eds.), *Communities, neighborhoods, and health: expanding the boundaries of place*, New York, NY: Springer, 2011. 21–34. doi: 10.1007/978-1-4419-7482-2_2

— *Space, place and health*. *Annals of GIS*, 21(2), 2015, 97–100. doi: 10.1080/19475683.2015.1007895

Harvey, Francis: *The social construction of geographical information systems*. *International Journal of Geographical Information Science*, 14(8), 2000, 711–713. doi: 10.1080/136588100750022741

Mark, David M and Turk, Andrew G: *Landscape categories in Yindjibarndi: ontology, environment, and language*. *Proceedings of the 6th International Conference on Spatial Information Theory (COSIT)*, 2003, 28–45. doi: 10.1007/978-3-540-39923-0_3

Massey, Doreen: *Living in Wythenshawe*. In: Borden, Iain; Kerr, Joe; Rendell, Jane; and Pivaro, Alicia (eds.), *The unknown city. Contesting architecture and social space*, Cambridge, MA: MIT Press, 2000. 458–474

Roche, Stéphane: *Geographic information science II: less space, more places in smart cities*. *Progress in Human Geography*, 40(4), 2016, 565–573. doi: 10.1177/0309132515586296

Smith, Barry: *On drawing lines on a map*. *Proceedings of the 2nd International Conference on Spatial Information Theory (COSIT)*, 1995, 475–484. doi: 10.1007/3-540-60392-1_31

— *Fiat objects*. *Topoi*, 20(2), 2001, 131–148. doi: 10.1023/A:1017948522031

Smith, Barry and Mark, David M: *Ontology and geographic kinds*. *Proceedings of the 8th International Symposium on Spatial Data Handling (SDH)*, 1998, 308–320

— *Do mountains exist? towards an ontology of landforms*. *Environment and Planning B*, 30(3), 2003, 411–427. doi: 10.1068/b12821

Smith, Barry and Varzi, Achille C: *Fiat and bona fide boundaries*. *Philosophical and Phenomenological Research*, 60(2), 2000, 401–420. doi: 10.2307/2653492

Varzi, Achille C: *Introduction*. *Topoi*, 20(2), 2001, 119–130. doi: 10.1023/A:1017944405193

Place and Placing Locations: A Cognitive Perspective

– Invited Keynote Paper –

Clare Davies 

Department of Psychology, University of Winchester, UK

Understanding and modelling places is an interdisciplinary problem, and one relevant but easily overlooked discipline is cognitive science. Many of the findings and intuitions that geographers and geographic information scientists have developed imply that places (at least, those that subtend a geographic area and do not have sharply defined boundaries) have a specific role and structure in human cognition: one of *categorizing* contiguous and semantically related locations, to optimize cognitive economy and efficiency. Thus “place”, in this sense, is a classification heuristic. This short paper will outline some of the new research questions that arise if we take this perspective on places, and suggest that computational and/or statistical models will need to be supplemented and “ground truthed” by human-participants work for useful progress to be made.

Keywords: classification; location; place cognition; semantic memory; research agenda

1 Introduction: Why Modelling Places Matters

Geographic information is ubiquitous and has increasingly become richer and more automatically updated. Modelling metric geographic space as objectively measured by science, however, can only take us so far. Our understanding of space in our human cognitive systems has many peculiar aspects that make it quite different from the space of a GIS, and the brain often does not seem that interested in accurately modelling space at all, preferring instead to prioritize what is visibly, semantically or emotionally significant (Davies and Peebles, 2010), and to simplify “uninteresting” aspects of the space between key vistas (Meilinger et al., 2014). Thus we know, from decades of research, that human spatial cognition closely links “what” and “where”, distorts distance and direction (and seems to record it non-transitively; Lloyd and Heivly 1987), and at the same time apparently imposes some kind of vague grouping and naming upon the space (Montello, 2003) to create (and usually to name) areas which we might think of as “places”.

Of course, “place” is more ambiguous and hence problematic as a term in English than “space” is. We talk loosely of “place” in smaller-scale spaces, in ways that are often synonymous with “location” (such as “my place at the table”, or “his place in the line”). We also use “place” to mean single functional buildings or locations in our environment: “come down to my place” (home); “that place on South Street” (shop, restaurant or bar); and “the place where he’s buried” (grave site). However, for the remainder of this paper, I will focus on the larger-sized meaning of “place” – an area of geographic space that is larger than one can see from a single point and thus is at least within the scale of what Montello called “environmental space” (Montello, 1993). Thus the focus will primarily be on urban or suburban localities – named but non-administrative “neighbourhoods” or districts within a city. The insights to be explored probably also apply to regions at the next scale up in Montello’s definitions, “geographic” space. Montello has also argued elsewhere that the geography of cognitive regions, as apparent groupings of locations in people’s minds, is distinctive from other types (Montello, 2003).

C Davies (2018): *Place and Placing Locations: A Cognitive Perspective*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 15–20

<https://doi.org/10.5281/zenodo.1472737>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

We already know, but it is worth restating, that understanding places of this kind is crucial for building a workable data model of urban (and often also of rural) geographic information. This, in turn, could greatly aid many organizations whose role forces them to liaise between formal spatial data and its associated professional expertise, and the messy, less easily predicted place-based geographies of the general public (Davies et al., 2009). Lives could be saved if ambulances avoided going to the wrong suburb or park. Location-based services would be far more accessible to users if intuitive notions of local place were included, rather than relying on formal addressing systems. Planners and military intelligence specialists would have a better understanding of public discourse, attitudes, and affiliations (the so-called “hearts and minds” knowledge) if they could model how a community divides and evaluates its local environment. These understandings might even allow all of these professionals to *predict* how people (en masse) might behave and move in crisis scenarios.

In 21st century society such professionals, unlike a hundred or even forty years ago, tend to be remote from the community they must support or protect and thus do not already share its understanding or knowledge. Place, then, is not a mere qualitative fancy for humanities scholars to muse about. Lack of understanding of it is costing lives and creating poorer-quality environments, here and now.

2 Geographic Information Science: Vague Vernacular Places

A key insight which GIScience has grappled with for some years is the notion that many places have vague, or fuzzy, boundaries. Web-sourced and other “big” data has allowed numerous demonstrations of this to be published in the past twenty years. Relatively few, however, have managed to check that the mappings they produced corresponded to human intuitions of the same places’ edges, rather than being artifacts of human error under particular circumstances (see Brindley et al. (2018) as a welcome exception). Underpinning much of this work appears to be an assumption that, if we capture enough geotagged data for a given area and solve the tricky problem of representing its vagueness within GIS, stable datasets of vaguely bounded places will eventually be produceable and usable.

So far, no work has established the speed with which vague boundaries may shift over time or even be contested between different subsets of the community in the first place, as is often been implied by much of the more qualitative human and social geographical research (Cresswell, 2014). Thus, major questions of quality, representativeness, timeliness, accuracy, and relevance are left unanswered. Whenever a new research study is published showing, typically, a kernel density model mapping some internet-sourced geotagged point data to establish vague place boundaries, at least ten research questions look beyond its findings:

1. Whose data does this represent, and which community members does it exclude?
2. Would the included community members be consistent about these boundaries in *other* situations?
3. When and why might people change their minds about placing a location within a named area?
4. How can we estimate a non-captured location’s probability of being in place X versus place Y?
5. What can we predict about the boundaries of a place for which we cannot gather enough data?
6. When do “hard” (crisp) boundaries apply instead – where a locality borders a highway or water-course? (Always? When does it spill beyond the linear feature, and why?)
7. Are all of the places we have modelled at the same hierarchical level? Are there other vague named regions which subsume or encompass them?
8. Are there any apparently *unnamed* places that people might also reference in ways missed by the usual data capture methods – e. g., localities referred to by a major street name? How do we identify and capture those?
9. How do people learn, decide upon, and perhaps evolve their shared communal knowledge of vague place extents?
10. Why does all of this happen in the first place?

The lack of theoretical grounding leaves us unable to answer a final, very basic, question about such work: is building a one-time dataset actually what we need to do? What if, instead, we need to generate predictive models specific to a given context and community, based on establishing certain parameters on an ad hoc basis? We may only be able to answer this when we understand better what feature of human cognition is producing the effect of vagueness and ambiguity in place understandings, what factors influence it, and the extent to which it depends on dynamically situated processes in a specific context rather than stable representations in memory. Thus we need to identify the fundamental psychological processes that create “place”.

3 Places as Semantic-Spatial Categories

Fuzziness is already a long-recognized feature of one particular area of human cognition: the concepts and categories we hold in semantic memory. Half a century or more of research in this field has established many often conflicting and unexpected aspects of how people choose to categorize objects and concepts into larger groupings (Murphy, 2002). The reason *why* they do so, however, is universally accepted: it is far more cognitively efficient to think of the world in terms of a smaller, organized set of concepts and types of object (or scenario, person, job, and so on) than to try to cope with the many thousands of individual items which we encounter over a lifetime (Bower, 1970). Thus categorization is part of the set of tools we use for heuristic – fast and simplified – cognitive reasoning and decision-making (Kahneman, 2011).

As I have pointed out elsewhere (Davies, 2009; Davies and Tenbrink, 2018), it makes sense to consider places in the same light. Grouping and naming an area of our city makes spatial problem-solving and language, and the retrieval of stored spatial knowledge relating locations together, far simpler and more efficient. Often, this simplification may come at the cost of precise metric spatial accuracy, but in many circumstances this does not actually matter. If I tell somebody that my grandmother lives in a given locality, it does not matter that their notion of that locality may be vague and different from mine, until they rely on the information to actually go there (at which point, we usually switch to more precise addressing notations). Where it *does* matter, as explained earlier, is where our human notions of place have to be interpreted by metric-space-only geographic information systems, and their less locally informed users.

Human-participants research, aiming to reapply some of the more complex insights about categorization to people’s local place knowledge, appears to support the basic claim that places are, mentally, a form of semantic category of locations, which happen to have spatial contiguity as a major (but by no means the only) dimension of similarity that links them together (Davies et al., 2018). There is also some suggestion from neuroscience that, although the two research domains almost never mention each other, semantic cognition and place knowledge (as a particular aspect of environmental-scale spatial cognition) are processed in contiguous and closely related areas of the human brain, specifically the anterior and medial temporal lobe (see, e. g., Lambon Ralph 2014; Lengen and Kistemann 2012). Other evidence suggests that place knowledge (especially of place names) gets damaged in semantic dementia along with recognition of objects and faces (Simmons and Martin, 2009; Snowden et al., 1994).

Thus we can tentatively conceive of places as categories that are partly spatial, but largely also semantic. Some fundamental insights that arise from this (based on insights from the semantic memory literature cited above) include:

1. Like concepts, places may be not just fuzzy at the edges, but show “graded membership” (often referred to as “typicality” – where every location may differ in the degree to which it is considered a good or typical exemplar of the place).
2. Most if not all places will have a common “core” area, which is less dependent on perceptual information and more semantically salient. (It will not necessarily be at or near the spatial centroid, however.)
3. There may be a degree of hierarchy, with larger places encompassing smaller ones, but some levels of the hierarchy may be privileged: for instance, one level (e. g., city) may be used in daily

life more often than the other levels, and people may be quicker to categorize a location into that level than into the smallest-scale level (e. g., neighbourhood).

4. Like categories, learning to identify a place may be gradual and incremental or may happen abruptly (e. g., from viewing a map).

Moreover, we have been able so far to show that places also conform to some of the less stable and challenging aspects of categories, investigated since the 1980s by cognitive scientists such as Barsalou (1985) and Hampton (2007). Thus, we have shown that the precise definition of a particular place may be sensitive to varying goals and contexts. Its boundaries (and the criteria used for judging them) may vary depending on the purpose and expertise of an individual thinker, and they can be influenced by cues from information sources such as maps (e. g., the precise cartographic placement of locality names).

This implies that to accurately represent places computationally, a stable spatial dataset may never completely suffice. Instead, we may need a dynamic, learned, context-adaptable *model*.

4 What Kind of Model?

Switching from talk of “mapping” to a requirement for partly semantic *classification* of locations into places raises a range of research questions. After all, at the time of writing Wikipedia was listing some 81 different algorithm types for classification of entities. Where to start?

First, we may consider the problem as one of clustering. If we took a hierarchical approach, should we take an agglomerative approach – locations get clustered together incrementally so that the number of divisions decreases over time? Or should we assume that the clusters are mostly already known – since childhood – and new residents in an area are likely to have heard of most local place names before they know exactly where they are? The latter insight would assume an approach analogous to partitional (“k-means”) clustering.

Second, individual locations are usually not independently “placed” (categorized); the placement of one location will influence the likelihood that the next remembered scene or landmark along its street will be similarly placed. However, the interdependencies are likely to be complex. Can we apply a “Dirichlet allocation”-style approach to model these? Similarly, where a boundary is “fuzzy”, which statistical distributions (probability curves) best represent that fuzziness? When is the slope gentler or steeper, or maybe even stepped? For example, perhaps sometimes the boundary between two localities will be conceived as the end of either one urban block or else the next; in more regular grid-pattern cities people’s assumptions may be less “fuzzy” than in other environments.

Third, places at the same granularity (e. g., urban localities or suburbs) appear to overlap in web-sourced data. Do they overlap in an individual’s mind too, or do people just assume that they are unclear about the boundary (but that there is one)? If overlap is accepted at some level, when does this happen (and not happen), and are people consciously aware of it – perhaps more so in some cultures or circumstances than others?

Fourth, supposing we build a place model by categorizing individual geotagged locations, as a number of studies have done in the past decade or so. How well does this reflect what people mean when they refer to that place as a whole (usually by uttering its toponym), rather than trying to classify locations into (or out of) it? In other words, how well is the concept or “essence” of the place reflected in the collection of locations that are probabilistically linked to it by such modelling? This is a question for qualitative, as well as quantitative, research.

Other questions relate to the details of how we categorize – which features and criteria we take into account other than the spatial contiguity of locations. In a particularly wealthy suburb, e. g., we may exclude a peripheral street because its housing units are smaller or of lower quality. A given position of a landmark within the street topology may sometimes matter more than its absolute spatial location, in deciding which place it “belongs” to. How far do the criteria vary with context and with which “crowd” (community) is being sampled? Can we abstract some general approximate “rules” or principles for a given type of geographic feature, so that criteria can be applied beyond spatial contiguity? These could help us to improve a machine learning algorithm trying to approximate locals’ understanding of place.

Finally, we need research to “ground truth” all such computational work. We have to ensure that the data we gather from “Big Data” or VGI sources, useful as it is, does correspond to the realities


of local people (and indeed, non-local visitors) for a given type of place. Some work already looks promising in this direction as mentioned earlier (see Brindley et al. 2018), but “where do you think you live?” is only one question among many which people have to consider about local places. Such work requires at least the three disciplines of psychology, geography, and computer science to work more closely together, possibly with additional insights from others, such as linguistics and sociology. There is plenty more place work to do.

5 References

Acknowledgements

Warm thanks go to the PLATIAL'18 organizers for inviting this contribution.

ORCID

Clare Davies  <https://orcid.org/0000-0003-0261-2353>

References

- Barsalou, Lawrence W: *Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 1985, 629–654. doi: 10.1037/0278-7393.11.1-4.629
- Bower, Gordon H: *Organizational factors in memory*. *Cognitive Psychology*, 1(1), 1970, 18–46. doi: 10.1016/0010-0285(70)90003-4
- Brindley, Paul; Goulding, James; and Wilson, Max L: *Generating vague neighbourhoods through data mining of passive web data*. *International Journal of Geographical Information Science*, 32(3), 2018, 498–523. doi: 10.1080/13658816.2017.1400549
- Cresswell, Tim: *Place. An introduction*. Hoboken, NJ, USA: Wiley & Sons, 2014
- Davies, Clare: *Are places concepts? familiarity and expertise effects in neighborhood cognition*. *Proceedings of the 9th International Conference on Spatial Information Theory (COSIT)*, 2009, 36–50. doi: 10.1007/978-3-642-03832-7_3
- Davies, Clare; Battye, Emilia; and Engelbrecht, Paula: *Places as fuzzy locational categories*, 2018. In preparation
- Davies, Clare; Holt, Ian; Green, Jenny; Harding, Jenny; and Diamond, Lucy: *User needs and implications for modelling vague named places*. *Spatial Cognition and Computation*, 9(3), 2009, 174–194. doi: 10.1080/13875860903121830
- Davies, Clare and Peebles, David: *Spaces or scenes: map-based orientation in urban environments*. *Spatial Cognition & Computation*, 10(2–3), 2010, 135–156. doi: 10.1080/13875861003759289
- Davies, Clare and Tenbrink, Thora: *Place as location categories: learning from language*. *Proceedings of Workshops and Posters at the 13th International Conference on Spatial Information Theory (COSIT 2017)*, 2018, 217–225. doi: 10.1007/978-3-319-63946-8_37
- Hampton, James A: *Typicality, graded membership, and vagueness*. *Cognitive Science*, 31(3), 2007, 355–384. doi: 10.1080/15326900701326402
- Kahneman, Daniel: *Thinking, fast and slow*. NYC, NY: Farrar, Straus and Giroux, 2011
- Lambon Ralph, Matthew A: *Neurocognitive insights on conceptual knowledge and its breakdown*. *Philosophical Transactions of the Royal Society B*, 369(1634), 2014, 20120392. doi: 10.1098/rstb.2012.0392
- Lengen, Charis and Kistemann, Thomas: *Sense of place and place identity: review of neuroscientific evidence*. *Health and Place*, 18(5), 2012, 1162–1171. doi: 10.1016/j.healthplace.2012.01.012

Lloyd, Robert and Heivly, Christopher: *Systematic distortions in urban cognitive maps*. *Annals of the Association of American Geographers*, 77(2), 1987, 191–207. doi: 10.1111/j.1467-8306.1987.tb00153.x

Meilinger, Tobias; Riecke, Bernhard E; and Bühlhoff, Heinrich H: *Local and global reference frames for environmental spaces*. *Quarterly Journal of Experimental Psychology*, 67(3), 2014, 542–569. doi: 10.1080/17470218.2013.821145

Montello, Daniel R: *Scale and multiple psychologies of space*. *Proceedings of the 1st International Conference on Spatial Information Theory (COSIT)*, 1993, 312–321. doi: 10.1007/3-540-57207-4_21

— *Regions in geography: process and content*. In: Duckham, Matt; Goodchild, Michael F; and Worboys, Michael F (eds.), *Foundations of geographic information science*, London, UK: Taylor & Francis, 2003. 173–189

Murphy, Gregory L: *The big book of concepts*. Cambridge, MA: MIT Press, 2002

Simmons, W Kyle and Martin, Alex: *The anterior temporal lobes and the functional architecture of semantic memory*. *Journal of the International Neuropsychological Society*, 15(5), 2009, 645–649. doi: 10.1017/S1355617709990348

Snowden, Julie; Griffiths, Helen; and Neary, David: *Semantic dementia: autobiographical contribution to preservation of meaning*. *Cognitive Neuropsychology*, 11(3), 1994, 265–288. doi: 10.1080/02643299408251976

Pinpointing Dream Settings onto Place Cookies

Cristina M Iosifescu Enescu  and Lorenz Hurni 

Institute of Cartography and Geoinformation, ETH Zurich, Switzerland

Dream reports are short pieces of text, where a dreamer summarizes the remembered experience of nightly dreams. Dream cartography addresses especially the spatial information contained in dream reports. In this context, the current formalization of space in GIScience such as points, lines, polygons, or labels, including place names or addresses, is not sufficient for mapping dream settings. In the best case, dream reports mention place names or streets. However, usually, the perception of space in dreams is designated in terms of whether this is familiar or not, inside or outside, safe or threatening. Moreover, basic comparisons between dream settings are meaningless with classic space definitions. This lead us to a different approach of space: the personal circles of places or, with a nickname, the place cookie. Here, the dream setting can be pinpointed at a certain distance from the centre of the cookie, which represents the familiarity of the setting to the dreamer.

Keywords: place cookie; dream cartography; familiarity of place; personal circles of places

1 Background

The research on dream cartography was initially envisioned to bring “new insights, through cartography, into the subject of dreams” (Iosifescu Enescu et al., 2015). However, we have discovered that also cartography and its fundamental concepts, such as representing space and distance, can profit from the insights developed for describing dream settings. Because dream settings are hard to fasten through traditional maps, we have researched other methods for describing the spatial dimension in dreams.

Dream settings are cases of platial data par excellence, as exemplified in Iosifescu Enescu and Hurni (2017). Although people were asked in structured questions about the countries, which appear in their dreams, when it came to open questions, they preferred to respond with dream places: e. g., childhood home, workplace, and holidays resort (Iosifescu Enescu and Hurni, 2017).

As mentioned above, dreams are very personal experiences and our scientific approach, which has the goal to abstract dream content in order to represent it visually, has to deal with many challenges regarding the diversity of the dream content or its description. Therefore, disassembling location not on its objective components, but on its subjective characteristics, on its qualities for an individual, serves the purpose of making places such as the dream settings comparable to each other. We consider the familiarity of a place to an individual, along with its time dependency, to be an eligible measure for settings in general, and not only for dream settings. The same works for social interactions in dreams. Abstracting the names of the persons appearing in one’s dreams and considering only their current relation to the dreamer makes the dreams comparable on social interactions.

In our project, being interested more in the “where” than in the “what” about dreams, we still cannot ignore the “what”, since in dream settings both aspects are highly intertwined. People recognize a place in a dream by certain elements. These can be elements from the natural environment, such as a river, a forest edge or a hill; or from the human-made structural elements such as buildings, specific landmarks, roads, or architectural style. People are also dreaming often about inside locations. Is

CM Iosifescu Enescu and L Hurni (2018): *Pinpointing Dream Settings onto Place Cookies*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18), pp. 21–27 <https://doi.org/10.5281/zenodo.1472739>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

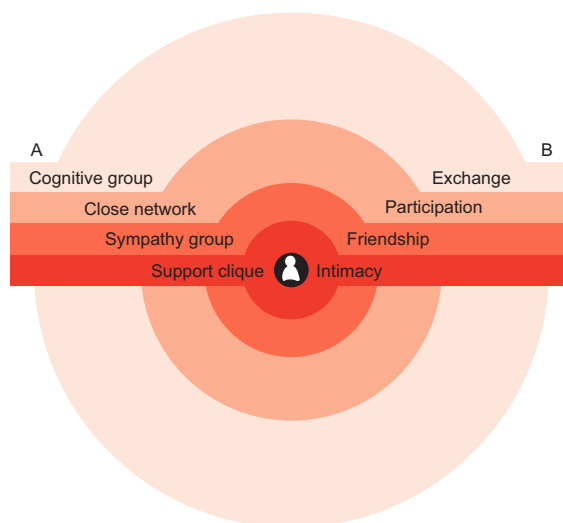


Figure 1: Concentric circles used for modelling the social network of a person. (A) Hierarchically inclusive levels of acquaintanceship by Dunbar (2010), (B) Circles of support by Falvey et al. (1997)

there the furniture, the wall decorations, the arrangement, and dimensions of the rooms, which betray the location? Actually, people are also in real life too little aware of the elements used to recognize a place. A dream setting might seem familiar to the dreamer, however, at times it turns out hard to identify why. Before dream cartography, dream research that considered settings only reported on whether this place was familiar or not, inside or outside, geographical or not (Domhoff, 1996; Strauch and Meier, 1996).

2 The Social Network

Research in sociology and anthropology (Dunbar, 1998, 2010; Hill and Dunbar, 2003) identified the size of the social network in humans to be in average 150 persons (the Dunbar's number), to which people maintain personal contact and this number is consistent to the group size predicted based on the size of the human neocortex (Dunbar, 1993). Moreover, four typical hierarchically inclusive levels of personal relationships were identified depending on the frequency of contact, therefore on the intensity of the relation individuals maintain with their peers.

The group, to which people have the closest contact (e. g., once a week) is the smallest in size and is called the support clique (Dunbar, 1998). It contains the closest friends or family, from whom one would seek advice, support or help. On the next level is the sympathy group, to which people maintain contact at least once a month (and which includes the support clique). Following is the band or the close network (which includes the previous two), to whom one makes a conscious effort to keep in contact. Finally, the cognitive group or the personal network is the biggest in size, and includes the other three groups (Figure 1A). Dunbar (2010) reveals also the expected size of these groups, in number of persons: roughly 5–15–50–150 (therefore increasing in size by a multiple of three). The classification can continue to the Supernetwork (500) or the Language Community (1500). In his book "How many friends does one person need?" Dunbar (2010) refers to these grouping levels as circles of acquaintanceship (Dunbar, 2010, p. 33).

Dunbar's idea of circles of acquaintances goes in the same direction as circle of friends, which has entered the common vocabulary. This other approach of circles in the social relationships evolved from a different perspective and is called the "circle of friends" or the "circle of support" (Falvey et al., 1997). The circle of support is used to create awareness and to actively enhance the number of people in an inner circle, e. g., in case of illness. Here, there are as well four concentric circles (Figure 1B). The innermost is the circle of intimacy and here are the people to whom one has intimate relationships: the partner, the close family. In the circle of friendship are people to whom one has regular contact, friends, and allies. The circle of participation is formed of people seen occasionally, with whom one

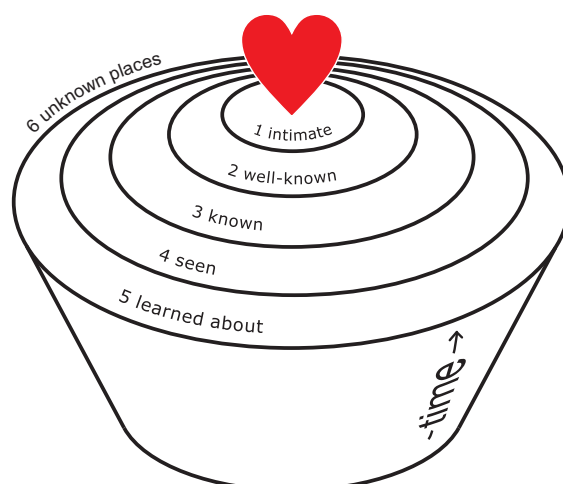


Figure 2: Personal circles of places (place cookie). Measuring places by their subjective distance to one's heart

shares interests: work colleagues, club members, or neighbours. The circle of exchange is concerned with paid relationships, such as doctors, teachers, therapists, cleaning personal, etc. It has been observed that people with disabilities tend to have a higher number of people in this last circle (paid professionals) than in the circle of friendship or in the circle of participation (Pearpoint, 1991). An important difference is that, whereas by Dunbar the number of persons contained in a circle is given in average, by Falvey et al. (1997) the number of persons in the two middle circles may be actively increased in order to achieve more support, become integrated into community and influence the personal wellbeing.

3 The Place Cookie

Analogous to the concentric circles of the social network, we propose the circles of intimacy in the knowledge of space, or the place cookie, nicknamed after its appearance (Figures 2 and 3), but also after the characteristic of saving personal relevant information (Web cookie). In the social relationships, Dunbar designates the innermost circle to have the higher, the outmost circle the lowest frequency of contact. This corresponds to a decrease in intimacy (most intense relations in the inner circle) and at the same time an increase in size towards the outer circle.

Applying this reasoning to places, we have the very familiar places in the inner circle and the less familiar places in the outer circles, outside remaining the unknown places. Examples of very familiar places are the current home, or childhood places – places to which the individual has close, intense relationships. Less familiar, but still well-known are, e. g., the work place, the way to work or to school, friends' homes, or other places people visit often. Next would be, e. g., holiday destinations (some people have a map with pinpoints on it) or other places, which have been seen only once, or did not occupy the memory, the attention for a long time. Places we learned about, but we have never been to, would be in the outer circle. Unknown places would be outside. To situate a place onto the personal circles of places (onto the place cookie), we propose an ordinal rating scale. The individuals will answer the following question, rating their familiarity/intimacy to a place on a scale from 1 (very familiar) to 6 (totally unfamiliar). For each number on this scale we also give a qualitative description (e. g., intimate, such as home) to make the rating easier and to make sure that the answers of different individuals can be compared in the end.

How do you rate your intimacy (your relationship) to this place?
(1 – very close to 6 – no relationship)

1. Intimate (such as home),
2. Well known (such as the work place or school class, the way to work or to school),
3. Known (places you often go to: a nearby park, the fitness studio, etc.),

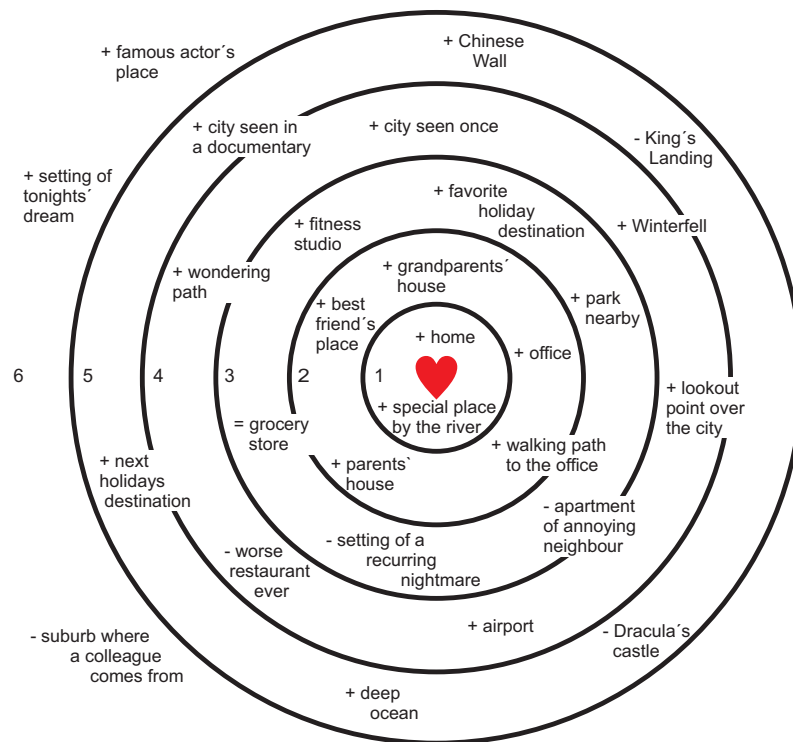


Figure 3: Example of a filled place cookie

4. Seen (such as a one-time holiday destination),
5. Learned about (read in a book, seen in pictures, maps, etc.),
6. Unknown – outside the concentric circles.

On the place cookie we can visualize the familiarity of a single place, but also the platial difference between two places. Yet the rated familiarity is an ordinal measure and the intervals between the circles are not necessarily equal. Therefore comparisons (e. g., a place is more familiar than another) are allowed; differences, however, are to be enjoyed with caution (e. g., the difference in familiarity is 2 points). To make it more comprehensive: this scale is similar to the school grades (both are of ordinal type).

Regarding the time dimension, we envision a deepness of the place cookie, the upper layer representing the current situation. With time, a place can undergo an upgrade in familiarity to a more inner circle, depending on the emotional intensity with which that place is being experienced; but this can happen also based on the changing function of that place over time (e. g., new workplace, new home). An old workplace, an old home, etc. gets a time stamp and moves into the deepness of the place cookie. Another example regarding a change in the function of a place to an individual is when an old home becomes the home of a friend. Here we can notice that a real, physical place can appear twice on the personal circles of places. In the former example there are two instances of the same place: one in the most inner circle, with a “past” time stamp, and one in its current situation. Moreover, the classification of places in the place cookie may include virtual, fictional places (imagined, watched movies, read books, explored on pictures or maps, experienced in virtual games and reality) and these can be part of any circle, depending on the relationship of the individual to them.

In the previous, merely positive relationships to a place were mentioned. The question remains open, if negative relationships to a place (e. g., a place where the person had a car accident or where he or she suffered mobbing, etc.) belong in the same circles as the positive ones – so only based on the intensity of the relationship and ignoring its valence. This would be analogous to having one’s enemies in the same circle as one’s friends. Making a trade-off between the valence and the intensity of the emotions (arousal) related to a place, we can keep it simple and consider only the intensity, the familiarity of a place to an individual. Figure 3 shows places that can be familiar (from 1 to 6) and



Figure 4: Map using the place cookie as symbol for spatial locations. Map © OpenStreetMap contributors (cf. www.openstreetmap.org/copyright)

in their valence positive (with +), negative (with –) but also neutral (with =) to a person. However, a study on virtual world-games (Hologatch et al., 2017) shows that the development of an enemy list has a different dynamic than a network of friends: characters with more enemies being more likely to attract other enemies, which does not hold for friendship or communication relationships. If other mechanisms are in place for remembering negative emotions relative to places, the simple solution could be to draw two place cookies: one for positive and one for negative emotions related to the place. The familiarity of places is then distributed on two place cookies.

4 Example of Application in Dream Cartography

A dream setting may be named by the place name, e. g., “Helvetiaplatz”, by its significance to the dreamer, e. g., “a place where I pass by with the bus every work-day and never stop”, or both at the same time. When this happens, we take advantage and place our place cookie with the interpreted familiarity (this would be, e. g., 4 in this case) on the city map. If the dreamer mentions her office and its location (e. g., ETH Zurich main building) in the same or in another dream, we can calculate the appearance of the place cookie for the office (e. g., 2) and place it on the same map at its spatial location. This results into a map (Figure 4), where both the spatial distance between the two places, and the platial difference of these places for the dreamer can be represented, in terms of familiarity, using the place cookies.

Another application of place cookies in dream cartography is in visualizing dream series. It allows, e. g., an aggregation of all dream settings in only one place cookie, showing the occurrence of places with different familiarity in a person’s dreams. It also makes possible to display the evolution of a place regarding its familiarity in a dream series. Moreover, comparisons of places in dreams between different individuals may be performed and are easier to visualize: e. g., the nightmares of the person A take place more often in unknown than in familiar locations, whereas those of the person B take place usually in very familiar locations.

5 Conclusions

To dream cartography, the classic distinction between space and place is relevant: the place has a specific meaning for a person, whereas space is general, with no implied reference to an individual. More precisely, the place is the space directly experienced by the people (Tuan, 1977). A taxonomy of

such places is given (Figure 2), where these are described with their qualities for an individual and, even if labeled (such as “Winterfell” or “Dracula’s Castle”) as shown in Figure 3, they imply a certain significance for the individual.

Using place cookies, dream settings become comparable. If a person dreams that she was at home and in the next dream scene in an unknown city, the difference, the distance between these two places on the place cookie is the biggest possible. If two dreamers dream being at the workplace, respective at the school, then the dream settings are quasi the same, situated on the same distance from the person’s heart, even if the mentioned workplace is in one country and the mentioned school in another. The pinpointing of dream settings over time on the place cookie is both easier to achieve and meaningful, making abstraction of sensitive (where is home) and at the same time data intensive information (at which time is the location how familiar; using only the date when the dream occurred). Therefore, if a person dreams about a virtual world after one evening of playing a virtual game, this points to a certain distance on the place cookie (e. g., 5 or less), but if it happens after a whole year of playing the same game, this distance decreases. Moreover, this virtual world place becomes comparable, on the place cookie, to a real place that appears in a dream of this person.

Similarly to the circles of support, through dream cartography we can notice, if in a dream series (reported dreams of one person in an uninterrupted interval of time) some inside circles are not enough occupied. On this bases we could recommend a more intensive contact and interaction to specific places and therefore support the upgrade of places in the place cookie from the position 4 (seen once) to 3 (known) or from 3 to 2 (well-known), and ultimately increase the wellbeing of an individual.

Funding

CM Iosifescu Enescu was supported by the Swiss National Science Foundation.

ORCID

Cristina M Iosifescu Enescu  <https://orcid.org/0000-0001-8127-2021>

Lorenz Hurni  <https://orcid.org/0000-0002-0453-8743>

References

- Domhoff, G William: *Finding meaning in dreams. A quantitative approach*. Plenum, 1996
- Dunbar, Robin IM: *Coevolution of neocortical size, group size and language in humans*. Behavioral and brain sciences, 16(4), 1993, 681–694. doi: 10.1017/S0140525X00032325
- *The social brain hypothesis*. Evolutionary Anthropology, 6(5), 1998, 178–190
- *How many friends does one person need? Dunbar’s number and other evolutionary quirks*. Cambridge, MA: Harvard University Press, 2010
- Falvey, Mary A; Forest, Marsha; Pearpoint, Jack; and Rosenberg, Richard: *All my life’s a circle. Using the tools: Circles, MAPS & PATH*. Toronto: Inlucision, 1997
- Hill, Russell A and Dunbar, Robin IM: *Social network size in humans*. Human nature, 14(1), 2003, 53–72. doi: 10.1007/s12110-003-1016-y
- Holovatch, Yuriy; Mryglod, Olesya; Szell, Michael; and Thurner, Stefan: *Analyses of a virtual world*. In: Kenna, Ralph; MacCarron, Máirín; and MacCarron, Pádraig (eds.), *Maths meets myths: quantitative approaches to ancient narratives*, Zurich: Springer, 2017. 115–130. doi: 10.1007/978-3-319-39445-9
- Iosifescu Enescu, Cristina M and Hurni, Lorenz: *Fictional volunteered geographic information in dream cartography*. International Journal of Cartography, 3(1), 2017, 76–87. doi: 10.1080/23729333.2017.1301627
- Iosifescu Enescu, Cristina M; Montangero, Jacques; and Hurni, Lorenz: *Toward dream cartography: mapping dream space and content*. Cartographica, 50(4), 2015, 224–237. doi: 10.3138/cart.50.4.3137

Pearpoint, Jack: *From behind the piano. The building of Judith Snow's unique circle of friends*. Toronto: Inclusion, 1991

Strauch, Inge and Meier, Barbara: *In search of dreams. Results of experimental dream research*. Albany, NY: State University of New York Press, 1996

Tuan, Yi-Fu: *Space and place. The perspective of experience*. Minneapolis, MN: University of Minnesota Press, 1977

The Near-Decomposability Paradigm Re-Interpreted for Place-Based GIS

Thomas Blaschke  and Sepideh Tavakkoli Piralilou 

Department of Geoinformatics – Z_GIS, University of Salzburg, Austria

We hypothesize that humans tend to think in objects while nature can be interpreted as gradients of matter and processes. Such gradients can be steep, like the borders between water and land or between forest and pasture. But objects and decomposing complexity call for a scale. For place-based GIS, the scale issue and the ability of handling multiple scales are even more crucial than for classic GIS. We argue that the paradigm of near-decomposability of systems can play an important role in GIScience research and for the foundation of place-based GIS.

Keywords: hierarchical patch dynamics paradigm; object-based image analysis; scale detection; Brownian motion; Wiener process

1 Introduction

1.1 Motivation

Geographic information systems (GISs) were originally designed to digitally represent physical entities, such as roads, buildings, parcels, or trees. Through increased modelling capacities, GISs were soon used to address invisible or not directly mappable information, such as suitability zones or potential maps. Over the years, we have witnessed frequent attempts of place-based investigations into human phenomena in the humanities and social sciences, but short literature searches on Google Scholar and Scopus reveal that this notion has only recently transgressed into geographic information science (GIScience), and only a small fraction of the GIS-relevant literature (most likely less than 0.1 per cent) deal with human-centred and philosophical notions of place.

The workshop Platial'18¹ calls for “platial” analyses. In previous work, the authors of this article have been criticized that “platial” is not a proper English word. We will, therefore, not use this term as such but will address the workshop questions accordingly. We will focus on what objects are and will link the GIScience notion of objects to the near-decomposability paradigm and to the hierarchical patch dynamics (HPD) paradigm, to find out if such methodologies can substantiate a multi-scale object centred methodology for place-based GIS.

1.2 Hierarchy Theory

According to Nobel prize laureate Herbert Simon, basically all viable systems, whether physical, social, biological, or artificial, have a near-decomposable architecture. They are organized into hierarchical layers of parts, sub-parts of parts, parts of sub-parts, and so on, in such a way that interactions between elements belonging to the same parts are much more common than interactions between elements belonging to different parts (Egidi and Marengo, 2004). To exemplify this notion, Simon uses the

T Blaschke and ST Piralilou (2018): *The Near-Decomposability Paradigm Re-Interpreted for Place-Based GIS*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 29–36

<https://doi.org/10.5281/zenodo.1472741>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

often-cited example of the two watchmakers who each have to build a watch out of 1000 pieces. One watchmaker assembles all pieces, one after the other, and has to start from scratch if he is disturbed in the process and forced to put the watch down, causing the pieces to fall apart. The other watchmaker designed his watches so that he could put together sub-assemblies of about ten components each, and each sub-assembly could be put down without falling apart. Ten of these sub-assemblies could be put together to make a larger sub-assembly, and ten of the larger sub-assemblies constituted the whole watch. This second, hierarchically structured strategy turns out to be more successful. Simon concludes that systems that are near-decomposable are much less vulnerable than systems that are not, as disturbances are more likely to remain confined to specific subcomponents. Near-decomposable systems limit interactions and information flows among different parts of the system and are, thus, better able to keep damaging events confined to sub-parts. While there is a large body of literature that utilizes, adopts, or criticizes the work of Simon and others in the fields of economics and operations research, there is surprisingly little evidence in GIScience literature. Here, instead of efficiency and stability of systems, the focus is on the spatial organization of systems.

In ecology and landscape ecology, hierarchy theory has been translated into a powerful way of dealing explicitly with spatial heterogeneity and has emerged as a unifying concept across different fields of earth sciences. Wu (1999) suggested the integration between hierarchy theory and patch dynamics through the emergence of the hierarchical patch dynamics paradigm (HPD) and laid a theoretical framework for a theory-driven break down of ecological complexity through a hierarchical scaling strategy. Most of the systems we examine in ecology and environmental science are characterized by organized complexity (Allen and Starr, 1982). On one hand, these systems have more components than analytical mathematics can handle; on the other hand, the use of traditional statistical methods cannot be justified because of the inadequate number, and the non-random behaviour, of components. Wu (1999), drawing on the concept of flux rates in hierarchy, suggests that ecological systems are near-complete-decomposable (or near-decomposable) systems because of their loose vertical and horizontal coupling in structure and function.

1.3 Object-Based Image Analysis

Blaschke (2010) summarized the utilization of multi-scale image segmentation methods in an attempt to bridge remote sensing and GIS functionality for the classification of remote sensing imagery while building on earlier work related to landscape ecology (Burnett and Blaschke, 2003). Image segmentation is a good example of how to spatially decompose complexity. Resulting objects at a certain scale are believed to have less internal heterogeneity concerning one or several parameters as compared to the adjacent areas. Tiede (2014) demonstrated that this concept can also be applied to non-image data. Despite its proven application to non-image data, this field is currently called (geographic) object based image analysis (OBIA/GEOBIA), see Blaschke et al. (2014).

1.4 Translating Concepts Between Disciplines

We want to transfer some elementary OBIA concepts to place-based GIS research. First, we need to expand the term “heterogeneity” from the original values created by sensors to encompass all kind of intensities of matter, processes, or interaction and their respective intensities. Here, by again referring to Simon, we may assume that for a more “intense” interaction the behaviour of one component depends more closely on the behaviour of other components belonging to the same part than on components belonging to other parts (i. e., objects vs neighbouring objects). This defines a near-decomposable system: individuals within a hierarchical subunit have closer, more widespread, more intense, and more frequent interactions than individuals belonging to different subunits. But a similar architecture can also be found in most complex artefacts, which consist of assembling parts and components that, in turn, can be assemblies of other parts and components, and so on, or in software (with the use of subroutines, particularly in object-oriented programming, Egidi and Marengo 2004).

2 Methodology

2.1 GIScience Objectification

Humans embrace natural complexity via abstraction. Our limited sensual and mental facilities cannot assimilate the whole. The challenge we face is how to leverage the strengths of abstraction by objectification (upon which so much of GI infrastructure is based) while learning from the concept of ecosystems as mobile expanses of flux, of complexes of material or energy gradients. We believe that decomposition allows us to grasp more of natural complexity. Burnett and Blaschke (2002) contrasted the nature of objects with gradients found in nature and addressed less anthropomorphically rendered “objects”, such as the margins of semi-natural forests and meadows. GIScience scholars deal with spatial objects but less often with gradients.

Without going into detail, we may refer to earlier work on the dichotomy of the vector and the raster domains in GIScience (see, e. g., Kemp (1996)). For simplicity’s sake, we may assume that space is most often manipulated, conceptually and in GI tool development, as complexes of objects – or as raster data. One efficacy of the object paradigm is obvious: one can *do* things with a partitioned space, e. g., a landscape that one cannot do with an un-partitioned space. Measuring the perimeter-to-area ratio of a forest patch, for instance, may provide insight into the nesting requirements of a particular bird species. However, as Burnett and Blaschke (2002) point out, there exists a paradox when many attributes used to define objects are derived explicitly from the placement of the object’s boundary itself. That raises the question of which comes first, the object or its boundary? If boundaries are not “real”, what alternatives do we have? While this is already difficult in the spatial domain – despite significant progress over the last fifteen years or so – it is even more ambiguous when dealing with place.

2.2 GIScience Partitioning of Space

Objectification requires spatial (and spatio-temporal) discretization. For a discussion on fiat vs bona fide objects, we refer to Smith and Varzi (2000). Any discretization of space may be likened to cutting with a double-edged blade. But dissection space is not an end in itself. Even post-dissection the partitioned space is doable today. In fact, in OBIA, objects are increasingly over-segmented to build complex objects from segments which may be regarded as object candidates (Burnett and Blaschke, 2003). It is well understood that in ecosystems, the structure or “gradients” that we observe are evocations of self-organized thermodynamically open systems (Müller, 1998). Goodchild et al. (1994) noted that there are many ways of representing a field as a collection of discrete objects. Objects in GI systems are simply the human discretization of near-decomposable hierarchical structures, and we lose sight of their “not-object” nature at our peril. The challenge is now to transfer these concepts to place-based (patial) GIS.

As stated at the beginning, in GIS we store information about the gradients that *define* apparent objects. This is made easier by the observation that gradients, though ubiquitous in nature, are not randomly distributed. We can draw upon the power of hierarchy theory (Koestler, 1967; O’Neill et al., 1986; Simon, 1969, 1977), a dialect of general systems theory (Bertalanffy, 1969), and identify holon (Janus-faced, near-decomposable elements) boundaries by characterizing the structure of the system at higher and lower levels.

It is difficult to classify hierarchies because their composite holons tend to overlap (Koestler, 1967). Koestler proposes a rough classification where structural hierarchies emphasize the spatial aspect of the system (a structure of the spatial domain or physical space itself), whereas functional hierarchies consider a process as another aspect. We seek a spatial partitioning schema that explicitly incorporates hierarchy and functional relationships. OBIA is somewhat successful in this respect (Blaschke et al., 2014). The artefacts that are manipulated in these software environments are, as we propose, closer to the holons found in nature; they are objects but can also be “not-objects” (Burnett and Blaschke, 2002).

2.3 Mathematical Foundations

Also, for place-based GIS, mathematical models are needed that can be adjusted to multiple scales. Especially at the nano and microscale, a discrete and stochastic description of processes is key, e. g., in

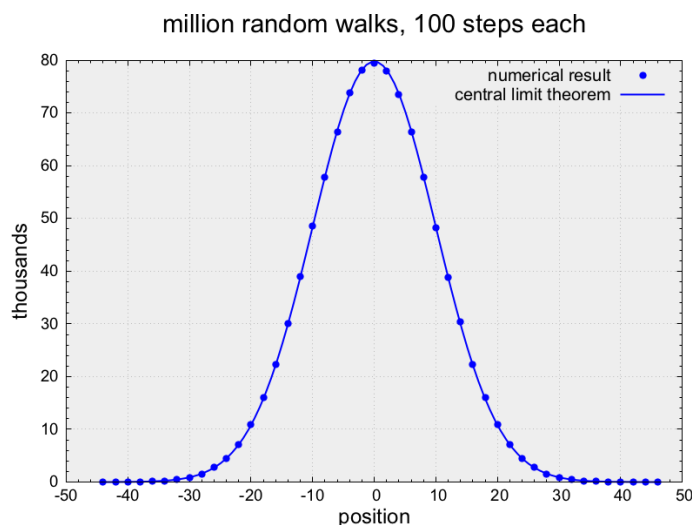


Figure 1: Central Limit

molecular and cellular biology or in nanophysics of materials and fluids. Because the dynamics and the properties of macroscopic systems must be understood, controlled, or even designed, the models on different scales have to be linked, and the information produced on the various levels has to be transferred. Finding transmission conditions in a computable way is an important goal for mathematics and computational sciences. We suggest to join analytic and computational bridging of scales, also taking into account hardware developments that can account for multi-scale and parallel structures. We argue that there are two principal ways of bridging scales analytically and computationally:

1. Symmetry concept. If different outcomes are equivalent, they should have the same probability.
2. Universality. Many random microscopic subsystems interact with each other. They may all look similar and may interact with each other to produce a larger outcome (an example is the central limit theorem).

Central Limit. The central limit theorem is a probability concept. The average distribution of a large number of identically distributed and independent variables is approximately normal. When certain conditions are fulfilled, and the number of iterations on independent variables is large enough, each of these variables will have a well-defined variance along with a well-defined expected value. Then, an approximately normally distributed function will result, as depicted in Figure 1.

Brownian Motion and Wiener Process. Brownian motion process, or *Brownian motion*, usually describes the physical process of movements originally observed by Robert Brown, who discovered the movement of tiny particles suspended in a liquid. It is a continuous-time stochastic process with independent, stationary increments and represents the motion of a point whose successive displacements are random and independent as well as statistically identical over different time intervals of the same length. Einstein considered the possibility that formalizing Brownian motion could support the idea that molecules existed. In mathematics, Brownian motion is described by the *Wiener process*, a continuous-time stochastic process named after Norbert Wiener. The Wiener process can be constructed as the scaling limit of a random walk or other discrete-time stochastic processes with stationary independent increments. Like the random walk, the Wiener process is recurrent in one or two dimensions (meaning that it returns almost surely to any fixed neighbourhood of the origin infinitely often). Unlike the random walk, it is scale invariant.

2.4 Utilizing Brownian Motion and the Wiener Process for Scale Detection

In computer vision, scientists developed multi-scale algorithms for the selection of salient regions (e. g., Kadir and Brady 2001). Such scale space-based approaches propose that saliency is defined in terms of local signal complexity based on the Shannon entropy of local image descriptors. The applicability of

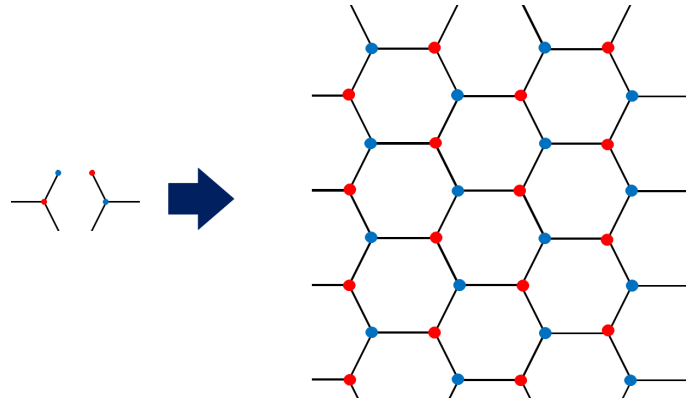


Figure 2: Two-dimensional analogue of random walk (Hairer, 2018)

this low-level approach has been demonstrated for several image processing problems. There are, of course, dozens of other scale selection frameworks in computer vision, some of which focus on saliency of points or pixels. Several computer vision approaches utilize a probabilistic localized scale selection principle, e. g., based on maximum likelihood estimation under a Brownian image model. Most of these methods aim to identify an “intrinsic” or “appropriate” scale for local image structures. Blaschke and Hay (2001) applied a *linear scale-space framework* from computer vision to automatically analyse real-world structures at multiple scales in a satellite image. They showed that when there is no a priori information about these structures, appropriate scale(s) for an analysis may be identified by applying Gaussian filters to an image at a range of kernel sizes resulting in a scale-space cube or “stack”, where each layer in the stack represents convolution at a specific scale. While some non-linearity situations could be identified through this “stack of scales”, the authors could not provide a mathematically sound foundation for the detection of such non-linearities.

We aim to transfer some basic and widely acknowledged methods from mathematics and from computer vision to scale detection problems in GIScience. The simulation of a two-dimensional random walk is presented in Figure 2. At a very low scale, a random function is assumed:

$$h: \text{Grid} \rightarrow \mathbb{Z} \quad \text{such that} \quad |h(x) - h(y)| = 1 \text{ for } x \sim y.$$

This means the next value is always either the previous value plus one or the previous value minus one. Alternatively, if one selects the value of any point, it will be assigned to a value from any existing neighbour and the difference will always be one. The random function results in a two-dimensional graph such as a honeycomb graph. The large amounts of the grids in Figure 2 results from universe object mechanism presented in Figure 3a. The simulation of an infinitely large grid is called a free field (Figure 3b), which is actually not a random function but a random distribution. As the free field is infinitely large, it should converge to something, i. e., a Gaussian generalized function following the equation

$$Eh(x) \cdot h(y) = -\log|x - y|,$$

where E denotes the expected value operator. But this is not proven yet. For the full paper, we shall verify whether the planar Gaussian free field can serve as a universal scaling limit for spatial scaling models, similarly to Brownian motion, which is the scaling limit of a wide range of discrete random walk models.

It is not possible to evaluate individual points in the free field. When looking up the image very closely one can see in the light coloured region a lot of dark spots and, likewise, dark regions contain light points, whereby the colour is just the value of the function. An increasing image resolution yields more and more points in the light region and the value increasingly fluctuates without converging to anything tangible. However, a value converges to a limited interval. Therefore, if we look at a little circle instead of the value of a point, we can find an average value. Then, this average value can converge to a limit by increasing the resolution or scale, respectively. We should, therefore, refer the free field to random distribution rather than to random function. The free field behaves like a function but cannot be evaluated at the level of individual points. But the free field is still an object that has a

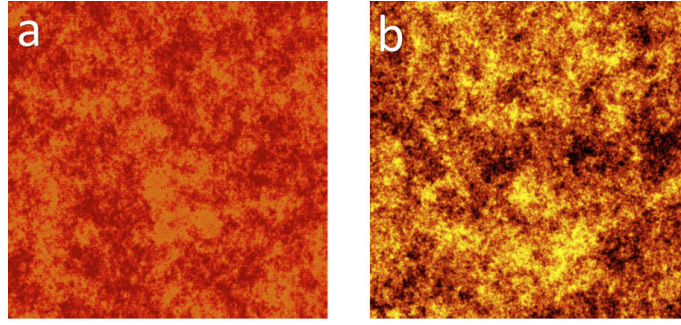


Figure 3: Simulations of Brownian Motion. (a) very large amounts of the grids, (b) infinitely large amounts of grids (free field)

description similar to the Brownian motion which is described by Gaussian distribution (Bell curve, see Figure 1). In the full version of this paper, we shall explain that this is relevant to place-based GIS for finding non-linearities of scaling and, under assumptions, defining appropriate scales.

2.5 Crossover Regimes

Let us take a “free model” according to the Gaussian universality classes and then perturb it a little bit. The models will look similar but if we zoom out, differences become visible. Therefore, bearing in mind the application for place-based GIS, what we can describe mathematically is the transition at which point to start a crossover from one universality class to another one. This transition can be approximated mathematically by normal form equations. We utilise (a) the Kardar-Parisi-Zhang (KPZ; Diehl et al. 2017) and, (b) the dynamic ϕ_d^4 model. These equations are used to investigate the dynamic scaling of growth processes (Sasamoto and Spohn, 2010). Recently, the KPZ equation has been considered to be a natural model for one-dimensional motion in the crossover regime (Hairer, 2015). However, the dynamic ϕ_d^4 model is for two or three dimensions. To explain large scale behaviours, both equations KPZ and ϕ_d^4 were experimentally evaluated for physical phenomena and, in particular, for the macroscopic behaviour of critical systems (Chandra and Weber, 2015). If we look at both equations in the right way, they can be considered as smooth. This means, when we measure the regularity, we are looking for how well we can approximate the function by polynomials (Hairer, 2015). As we restricted the KPZ to be one-dimensional, the spatial variable x can only take on values from a one-dimensional space (Chandra and Weber, 2015). Moreover, h is our random function and ϵ represents space-time white noise, which is a quite irregular random distribution. Thus, the KPZ equation can be defined as:

$$\partial_t = \partial_x^2 h + (\partial_x h)^2 + \epsilon \quad (kpz, d = 1).$$

For the dynamic ϕ_d^4 model, the spatial variable x takes values in two or three-dimensional spaces. The ϵ is the same as in the KPZ. The Gaussian free field can be assumed as a Gaussian random field on $\phi: R^d \rightarrow R$, which is actually a distribution and not a function. For the dynamic ϕ_d^4 model we can then assume

$$\partial_t \phi = \Delta \phi + c_1 \phi - c_2 \phi^3 + \epsilon \quad (\phi_d^4, d = 2, 3)$$

3 Conclusion and Outlook


We hypothesized – and will explicate further in the full version of this manuscript – that humans are programmed to grasp complexity through objects. We have referred to a particular GIScience methodology (OBIA) to decompose complexity as well as to the near-decomposability paradigm and some mathematical concepts. We conclude that GIScience needs to incorporate mathematical models to develop its theory and methodology further as exemplified with multi-scale handling. Still, some of the briefly illustrated concepts, in particular, two-dimensional fractional Brownian motion, are difficult to understand and to be utilized, both conceptually and computationally.


We believe that OBIA can serve as a methodology for place-based GIS, particularly when taking it out of the image processing domain, as illustrated by Tiede (2014) for an alternative overlay methodology. OBIA is less dependent on segmentation as one may think. There is never a perfect solution to segmentation. We argue to implement a flexible, yet theory-driven approach to build image objects on demand based on image primitives (segments). We suggest that generating a structural hierarchy needs to be followed by building a corresponding object-relation hierarchy. Some of such hierarchies can be derived directly from multi-scale image analysis (semantic rules can be created using sub/super-object spectral and spatial information such as neighbourhood, shape, size, compactness, etc.), but others must originate from expert knowledge, machine learning, or mathematical models. Expert knowledge will undoubtedly be incomplete; however, it allows us to partition space based on a spatial semantic network. By using fuzzy rules, we can deal with transition zones or gradients and move towards the incorporation of place as rule-based multi-scale objects and their relationships.

Notes

1. <http://platial18.platialsience.net>

ORCID

Thomas Blaschke  <https://orcid.org/0000-0002-1860-8458>

Sepideh Tavakkoli Piralilou  <https://orcid.org/0000-0002-1188-8290>

References

- Allen, Timothy FH and Starr, Thomas B: *Hierarchy. Perspectives for ecological complexity*. Chicago, IL: University of Chicago Press, 1982
- Bertalanffy, Ludwig von: *General system theory: foundations, development, applications*. New York, NY: Braziller, 1969
- Blaschke, Thomas: *Object based image analysis for remote sensing*. ISPRS Journal of Photogrammetry and Remote Sensing, 65(1), 2010, 2–16. doi: [oi.org/10.1016/j.isprsjprs.2009.06.004](https://doi.org/10.1016/j.isprsjprs.2009.06.004)
- Blaschke, Thomas and Hay, Geoffrey J: *Object-oriented image analysis and scale-space: theory and methods for modeling and evaluating multiscale landscape structure*. International Archives of Photogrammetry and Remote Sensing, 34(4), 2001, 22–29
- Blaschke, Thomas; Hay, Geoffrey J; Kelly, Maggi; et al.: *Geographic object-based image analysis – towards a new paradigm*. ISPRS Journal of Photogrammetry and Remote Sensing, 87, 2014, 180–191. doi: [10.1016/j.isprsjprs.2013.09.014](https://doi.org/10.1016/j.isprsjprs.2013.09.014)
- Burnett, Charles and Blaschke, Thomas: *Objects /not-objects and near-decomposability: ecosystems and GI*. Proceedings of the 2nd international Conference on Geographic Information Science (GIScience), 2002, 225–229
- *A multi-scale segmentation / object relationship modelling methodology for landscape analysis*. Ecological Modelling, 168(3), 2003, 233–249. doi: [10.1016/S0304-3800\(03\)00139-X](https://doi.org/10.1016/S0304-3800(03)00139-X)
- Chandra, Ajay and Weber, Hendrik: *Stochastic PDEs, regularity structures, and interacting particle systems*. arxiv:1508.03616 [math.AP], 2015
- Diehl, Joscha; Gubinelli, Massimiliano; and Perkowski, Nicolas: *The Kardar–Parisi–Zhang equation as scaling limit of weakly asymmetric interacting Brownian motions*. Communications in Mathematical Physics, 354(2), 2017, 549–589. doi: [10.1007/s00220-017-2918-6](https://doi.org/10.1007/s00220-017-2918-6)
- Egidi, Massimo and Marengo, Luigi: *Near-decomposability, organization, and evolution: some notes on Herbert Simon's contribution*. In: Augier, Mie and March, James G (eds.), *Models of a man. Essays in memory of Herbert A. Simon*, Cambridge, MA: MIT Press, 2004. 335–350

Goodchild, M; Buttenfield, B; and Wood, J: *On introduction to visualizing data validity*. In: Hearnshaw, Hilary M and Unwin, David J (eds.), *Visualization in Geographical Information Systems*, Chichester, UK: Wiley & Sons, 1994. 141–149

Hairer, Martin: *Bridging scales from microscopic dynamics to macroscopic laws*. Invited talk, 20 March 2018

— *Regularity structures and the dynamical ϕ_3^4 model*. arxiv:1508.05261 [math.PR], 2015

Kadir, Timor and Brady, Michael: *Saliency, scale and image description*. *International Journal of Computer Vision*, 45(2), 2001, 83–105. doi: 10.1023/A:1012460413855

Kemp, Karen K: *Fields as a framework for integrating GIS and environmental process models. Part 1: representing spatial continuity*. *Transactions in GIS*, 1(3), 1996, 219–234. doi: 10.1111/j.1467-9671.1996.tb00046.x

Koestler, Arthur: *The ghost in the machine*. London, UK: Hutchinson, 1967

Müller, Felix: *Gradients in ecological systems*. *Ecological Modelling*, 108(1–3), 1998, 3–21. doi: 10.1016/S0304-3800(98)00015-5

O'Neill, Robert V; DeAngelis, Donald Lee; Waide, Jack B; and Allen, Timothy FH: *A hierarchical concept of ecosystems*. Princeton, NJ: Princeton University Press, 1986

Sasamoto, Tomohiro and Spohn, Herbert: *One-dimensional Kardar-Parisi-Zhang equation: an exact solution and its universality*. *Physical Review Letters*, 104(23), 2010, 230602. doi: 10.1103/PhysRevLett.104.230602

Simon, Herbert A: *The sciences of the artificial*. Cambridge, MA: MIT Press, 1969

— *The organization of complex systems*. In: *Models of Discovery*, Dordrecht, The Netherlands: Springer, 1977. 245–261. doi: 10.1007/978-94-010-9521-1_14

Smith, Barry and Varzi, Achille C: *Fiat and bona fide boundaries*. *Philosophical and Phenomenological Research*, 60(2), 2000, 401–420. doi: 10.2307/2653492

Tiede, Dirk: *A new geospatial overlay method for the analysis and visualization of spatial change patterns using object-oriented data modeling concepts*. *Cartography and Geographic Information Science*, 41(3), 2014, 227–234. doi: 10.1080/15230406.2014.901900

Wu, Jianguo: *Hierarchy and scaling: extrapolating information along a scaling ladder*. *Canadian Journal of Remote Sensing*, 25(4), 1999, 367–380. doi: 10.1080/07038992.1999.10874736

Turin's Foodscapes: Exploring Places of Food Consumption Through the Prism of Social Practice Theory

Alessia Calafiore, Guido Boella, Elena Grassi, and Claudio Schifanella

Department of Computer Science, University of Torino, Italy

This contribution wishes to propose an addition to the existing toolbox of techniques employed to approach and render explicit the place semantics embedded in geosocial data. Inspired by the notion of relational place introduced by human geographers, we focused on people's experience of the city derived from the aggregation of the points of view of different social groups. We analysed socio-spatial behaviour under the frame of social practice theories, defining social practices as collective social actions performed by groups of people that display a similar behaviour. Applying spatial pattern analysis and clustering on data extracted from TripAdvisor platform, we classified social groups of users depending on our prior knowledge and their spatial behaviours.

Keywords: place; social practice; food consumption; TripAdvisor

1 Introduction

The wealth of information generated by users' interaction on social media and web-based platforms have been the object of extensive exploration in the last decade. A large part of the information collected via social media is nowadays georeferenced and this has brought such platforms to be considered valuable sources for geospatial analysis and urban planning, along with official data provided by institutional agencies (Ballatore and De Sabbata, 2018; Bello-Orgaz et al., 2016; Campagna, 2016; Guo et al., 2017; Kelley, 2013; Rzeszewski, 2018; Shelton et al., 2015). Data that populate geosocial platforms are produced by a non-expert public and have the potential to provide a description of the city derived from the aggregation of many heterogeneous points of view and experiences (Elwood, 2008). Therefore, insights that can be extracted from such sources should be interpreted under a *platial* rather than purely spatial perspective. Place identity, differently from space, is directly related to people socio-spatial behaviours. This contribution wishes to explore the socio-spatial behaviour that emerges from these new bottom-up sources of information under the prism of social practice theory. The concept of social practice has been considered by philosophers and social scientists such as Tuomela (2002), Schatzki (1996), and Giddens (1986) as the minimum unit to analyse people behaviour and describe social phenomena. Furthermore, social practices have been treated in urban studies (Brenner and Schmid, 2015; McFarlane, 2011)¹ and social geography (Massey, 1994; Murdoch, 2005; Soja, 1989) as the proxy to understand the multiple identities of urban places from a relational perspective.

The data employed in this analysis come from the TripAdvisor website², and are instrumental to test methods for clustering users' behaviour in terms of social practices. In particular we try to discover place-related regularities in the practice of food consumption in the city of Turin, Italy.

A Calafiore, G Boella, E Grassi, and C Schifanella (2018): *Turin's Foodscapes: Exploring Places of Food Consumption Through the Prism of Social Practice Theory*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18), pp. 37–43

<https://doi.org/10.5281/zenodo.1472743>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.



Figure 1: A typical restaurant page. Fields extracted are underlined in light grey.

2 A Platial Perspective on Cityscapes Through the Lens of Social Practice Theory

Place, differently from space, has been conceptualized as directly dependent on people experiences. As a consequence, our analysis aims at rendering explicit the social nature of people’s spatial behaviour. In order to do so, employing crowdsourced geographic data, we refer to social practice theory. There are many variants of social practice theory, among them the most comprehensive frameworks can be found in the works of Shove et al. (2012), focused on analysing changes in social practices and incorporating materiality in their conceptualization; Reckwitz (2002) approaching social practice theory from a culturalist perspective, therefore placing the social in relation with symbolic and cognitive structures of knowledge; Schatzki (1996) defining the site of the social as bundles of practices and material arrangements; and Tuomela (2002) considering social practice as a collective social action centred on the *we-attitude* of its participants. To the purposes of our work, we use Tuomela’s definition of social practices (Tuomela, 2002) since it specifically introduces the notion of *we-attitude* to unify participants of a practice by their shared intentionality. He claimed that “the core sense of a social practice is to be a repeatedly performed collective social action, because of a certain shared *we-attitude*, where the *we-attitude* must be a primary reason for the repeated activity, one without which the agents would not take part in it”. Therefore, given Tuomela’s conceptualization, social practices must be associated with a *we-attitude* which is shared within a group of people. For example, someone’s decision to go to a restaurant in the city centre can be occasional or recurrent – in the latter, we can assume that there is, besides the desire to go out for dinner or lunch, something else motivating the choice, e. g., that the person likes the city centre more than other areas. However, when we recognize a social trend, i. e., when the action is recurrent not only for an individual but for a significant number of people, we can consider it as a social practice, since what motivates the action is shared collectively (i. e., eating out in the city centre is considered to be “cool”). To identify this collective intentionality (*we-attitude*) we followed two directions. First, we selected activities performed by members of the same social group, which we knew a-priori from the datasets (Italian tourists, Turin locals, foreign tourists) and verified whether a convergence existed in their spatial behaviours. Secondly, we tried to detect emergent social groups, to cluster users according to their spatial behaviours, and to measure the similarity among cluster’s members and dissimilarity with others.

3 Datasets description

The information extraction of the TripAdvisor platform produced two datasets about Turin’s restaurants and the respective reviews. The first dataset consists of a scrape of the result page of the query: “restaurant in Turin”. Dataset one consists of 2116 observations, of which 1886 are within the Turin city boundaries. The addresses of all restaurants have been geocoded and mapped, and to each restaurant has been attached information about the cuisine type, ratings and average cost. Figure 1 shows a typical restaurant page in TripAdvisor highlighting the fields which have been extracted.

The second dataset consists of an extraction of each restaurant’s individual page, scraping reviews posted in four different languages: Italian, English, French, and Spanish. We count a total of 238,394 reviews from 2007 to 2016, of which 95% written in Italian and only 5% in the other languages.

Each review has a date, a title, and an excerpt, it is associated with the user id, user name, and the city where the user comes from. Not all user profiles have their origin set. We assume that all the reviews made in languages other than Italian have been added by foreign tourists. In order to distinguish among Italian tourists and Turin’s residents we subset the original dataset to have only users with the origin set and we filtered users whose origin is Turin or not (details on the distribution of total review by group are presented in Section 4.1). To count users we used as identifier the “userID”,

Table 1: Social groups on TripAdvisor

Social Group	Number of users	Number of reviews	Avg. reviews per user
Italian Tourists	45,404	93,961	2.069
Foreign Tourists	15,863	15,954	1.005
Turin Locals	28,000	104,815	3.743

“userName”, and “userOrigin”. However, some of them could not be recognized as individual users. This problem raises from the integration between Facebook’s login and TripAdvisor’s login. In some cases, indeed, we only know that the user is “A member of TripAdvisor on Facebook”. Depending on the privacy settings users can decide not to provide any information to TripAdvisor. These users cannot be clearly distinguished among each other, therefore they are not included.

4 Results

4.1 Comparing Social Groups’ Spatial Behaviours

In the present work we consider the activity of reviewing restaurants a proxy for two different social practices: that of using TripAdvisor itself and, assuming that users have actually visited the restaurants they review, the practice of eating out in Turin.

In light of the latter, we now try to identify different social groups in the data and compare their collective preferences regarding where to eat out in the city of Turin. We use our a-priori knowledge of users’ origins to filter reviews made by each social group: Italian tourist, foreign tourist, and Turin locals. The reviews from which we extract the users’ origin are 214,730 (90% of the total). Table 1 shows the number of users belonging to different social groups as well as the total reviews each group has posted. Although the number of reviews posted by tourists (Italian plus foreign) is slightly higher than the number of reviews posted by Turin locals, the latter are the most active in producing contents, posting 3.7 reviews on average. This is rather expected given that locals have more opportunities to visit Turin’s restaurants while tourists may be in Turin for a limited time and may not visit the city again.

To evaluate how reviews posted by different social groups are spatialized we employed a grid of 971 hexagons of 33,000 m² each³. The area has been chosen to approximate the average area of the 92 census tracts of Turin. The reviews posted by members of each social group have been aggregated by cell. The resulting spatial distributions have been transformed dividing by the maximum of each distribution to render them comparable. Figures 2a, 2b and 2c show the spatial distributions of the three social groups.

4.2 Identifying Emergent Social Groups from Spatial Behaviour

Now we try to identify *emergent* social groups of users, unified by a shared we-attitude towards eating out in certain areas rather than others. We specifically investigated user preferences in visiting certain neighbourhoods of the city. We employed the neighbourhood as geographical unit since it better refers to what Egenhofer and Mark (1995) define the “naive geography” of the city. Neighbourhoods have been defined as a “key living space [...] which symbolizes aspects of the identity of those living there to themselves and to outsiders” (Meegan and Mitchell, 2001). Given the strong symbolic meanings people associate with neighbourhoods, we hypothesized that identifying clusters of users, depending on their neighbourhood preference structure, may help us recognizing possible shared additional reasons (a we-attitude) that motivate people in performing the “eating out” activity in certain areas.

We applied the K-Means algorithm to the data, a commonly used, simple but generally rather efficient clustering method. Essentially, the algorithm intends to partition n objects into k clusters in which each object belongs to the cluster with the nearest mean. This method produces exactly k different clusters of greatest possible distinction. The problem of this method is that the best number of clusters k , leading to the greatest dissimilarity between clusters, must be decided a priori. Since the objective of K-Means clustering is to minimize total intra-cluster variance or the squared error

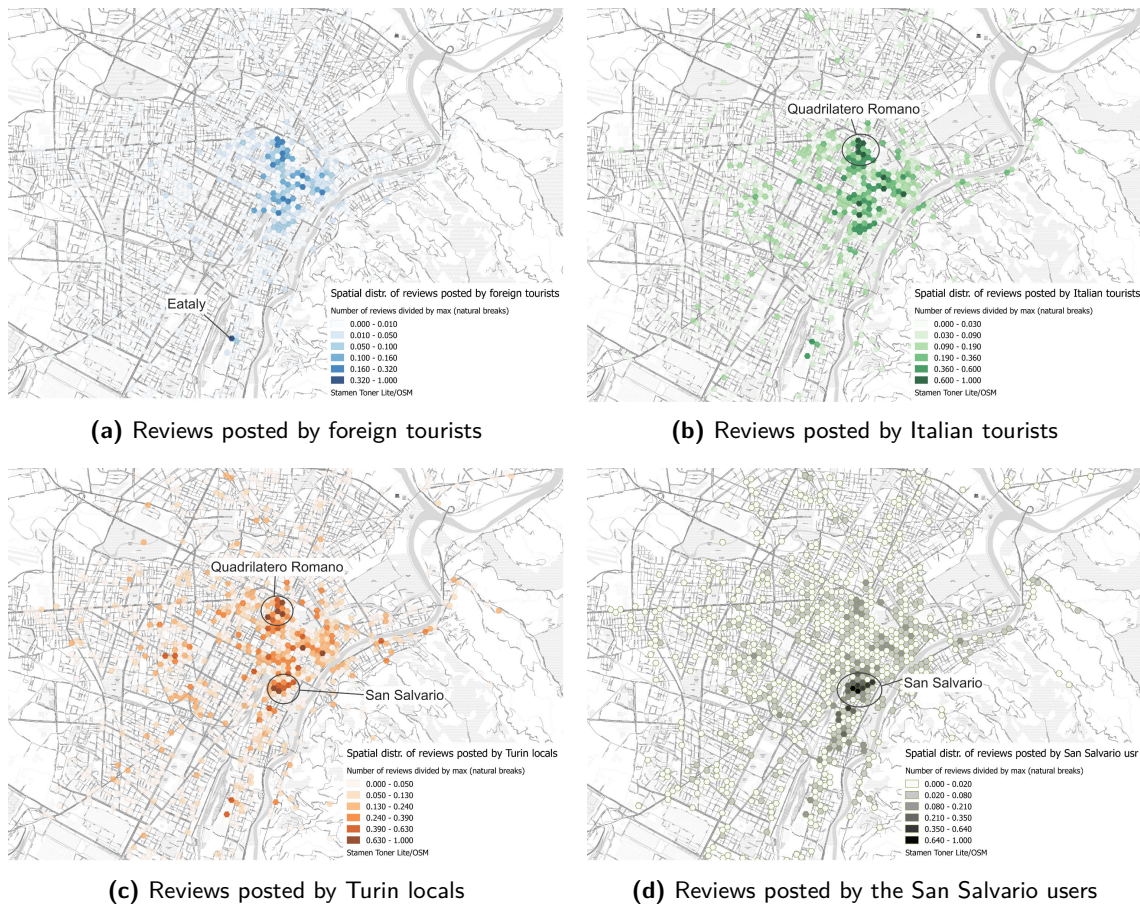


Figure 2: Maps displaying the popular areas for each social group. Spatial distributions have been transformed dividing by the maximum.

function, we can evaluate which number of clusters minimizes the squared error running it n times and look at the total within clusters sum of squares.

The errors decrease significantly until k equals 5, therefore, we run the algorithm to result in 5 clusters. We run the K-means algorithm set with $k = 5$ on an $m \times n$ matrix where m is the number of users and n is the number of Turin's neighbourhoods. Each a_{mn} corresponds to the sum of reviews the m -th user has posted in restaurants located in the n -th neighbourhood. It resulted that user groups display very similar behaviour regarding the favourite neighbourhood – the global maximum corresponds to the city centre in clusters 1, 2, 3, and 5. While the strong preference of dining out in the city centre is rather expected, the distribution of reviews made by users belonging to cluster 4 changes significantly. The latter, compared to other clusters, has a very low number of reviews in the city centre and a global maximum corresponding to the San Salvario neighbourhood. What the analysis seems to highlight is the existence of a group of users who consistently visit San Salvario but very seldom dine out in other areas of the city. No other neighbourhood, including the most visited part of the city, the city centre, shows such a pattern. In our view, this circumstance is particularly significant. As a consequence of this result, we decided to consider those users as members of an emergent social collective, performing the same social practice of dining out in San Salvario, which we called *San Salvario users*.

5 Mapping Turin Foodscapes

Figure 2 displays the distribution of tourists' (foreign and Italian), Turin locals, and San Salvario users reviews⁴. Foreign tourists, as expected, tend to visit mostly the city centre. However, the maximum

number of reviews posted by foreign tourists corresponds to the area where the Eataly restaurant is located (no other restaurants are present within that cell), in the southern district of Lingotto. Eataly is a famous multinational chain of Italian high quality food; the same place is similarly very much reviewed by Italian tourists (see Figure 2b), but less by Turin locals (see Figure 2c). Comparing to foreign tourists, the spatial distribution of reviews posted by Italian tourists shows a more varied distribution of the peak areas (dark green). A cluster of cells with the highest values of reviews corresponds to the Quadrilatero Romano area (a portion of the city centre that is also the oldest part of Turin). Figure 2c shows Turin locals reviews' spatial distribution, which is similar to that of Italian tourist but differs from it in two regards: Eataly is not in the highest classes and a cluster in the San Salvario area is visible in addition to the one in Quadrilatero Romano. From this analysis we recognized at least three social practices related to social groups: tourists (particularly foreign) visit Eataly, Italian tourists prefer to dine out in the Quadrilatero Romano area, and Turin locals have two hot spots, Quadrilatero Romano and San Salvario. A fourth social group was detected through the clustering as the one of San Salvario (Figure 2d), *aficionados*, which clearly shows a peak in the area of San Salvario and low usage of other areas.

The construction of an identity of San Salvario as the neighbourhood of choice for night life venues is not new and is confirmed also by other studies about the city, which described some of the existing and potential consequences of this phenomenon. In particular, the increase in popularity of the neighbourhood has impacted the life of the residents, divided between those who desire quiet nights and those who enjoy a lively neighbourhood and, as warned by Semi (2015), face the risk of a gentrification driven by the nightlife change⁵. Therefore, what resulted from the data, even if in relation to a single social practice (dining out), mirrors quite well a specific social perspective from which we can see a part of the city, particularly from the perspective of Turin locals and a mixed group which favours to dine out only in San Salvario.

6 Conclusions

This work explored methods to detect social practices related to food consumption in the city of Turin from crowdsourced data. The data sample we have extracted is about Turin's restaurants, so that the great majority of the reviews are written in Italian. Also, we assumed that no Italian resident has written reviews in languages other than Italian. Moreover, a clear bias of our study is given by the fact that reviews do not correspond to actual visits. TripAdvisor, as well as other platforms⁶, are now allowing to make table reservations. Such data may allow less biased analysis. Framing our work in social practice theories drove us to associate spatial behaviours with specific groups of people, which show a shared collective intentionality. To this regard we believe the approach proposed has the potential to render explicit the multiplicity which emerges from the different ways people experience and use spaces. As Mela et al. (2014) maintains, "the choice of a specific space implies a broader selective act, whether conscious or not: in accepting the possibility of encountering a particular set of individuals, with whom one identifies, also diminishes the probability of encountering another". The methods applied in this work are specifically designed to identify the selective function that people embed when choosing where to dine out, looking at behaviours that are similar within a group and dissimilar to the others. Our contribution shows that, even with a single source of information, it is possible to identify emergent behavioural patterns associated with specific social groups, which use the city in different ways. Further case studies with richer and diverse data sources are certainly needed. The approach presented here can, in fact, be applied to domains other than that of dining habits: possible applications include extracting knowledge about the use of public space through the recognition of social practices associated to different demographics (who visits public spaces, e. g., young/old, male/female geographies) or interests (why visiting certain public spaces, e. g., sport, culture, tourism).

Notes

1. Note that the two authors have different theoretical approaches to define urbanization process. However, in both cases urban practices play a crucial role in the characterization of urban spaces.

2. <http://www.tripadvisor.com>. The domain extensions employed to this analysis are .com to retrieve reviews in English; .it, for the reviews written in Italian; and .es and .fr, for the reviews in Spanish and French respectively.
3. We used hexagonal cells in line with recent works (Poorthuis and Zook, 2014; Rzeszewski, 2018; Shelton et al., 2014), which favour them to rectangular grids for essentially two reasons: hexagons can be more easily varied in size to address the modifiable areal unit problem and they share six instead of four neighbourhoods, which is an advantage for statistical analysis.
4. The number of reviews have been normalized dividing by the maximum of each distribution. Maps classes have been computed using the natural break function.
5. A similar pattern of change involved the Quadrilatero area, where the increasing number of restaurants and pubs was also accompanied by a change of the residents' social background.
6. for example, Open Table (<https://www.opentable.com>) and Yelp (<https://www.yelpreservations.com>)

References

- Ballatore, Andrea and De Sabbata, Stefano: *Charting the geographies of crowdsourced information in Greater London*. Proceedings of the 21st AGILE Conference on Geographic Information Science, 2018, 149–168. doi: 10.1007/978-3-319-78208-9_8
- Bello-Orgaz, Gema; Jung, Jason J; and Camacho, David: *Social big data: recent achievements and new challenges*. Information Fusion, 28, 2016, 45–59. doi: 10.1016/j.inffus.2015.08.005
- Brenner, Neil and Schmid, Christian: *Towards a new epistemology of the urban?* City, 19(2–3), 2015, 151–182. doi: 10.1080/13604813.2015.1014712
- Campagna, Michele: *Social media geographic information: why social is special when it goes spatial?* In: Capineri, C; Haklay, M; Huang, H; et al. (eds.), *European Handbook of Crowdsourced Geographic Information*, London, UK: Ubiquity, 2016. 45–54. doi: 10.5334/bax.d.
- Egenhofer, Max J and Mark, David M: *Naive geography*. Proceedings of the International Conference GIS, From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning, 1995, 1–15. doi: 10.1007/3-540-60392-1_1
- Elwood, Sarah: *Volunteered geographic information: key questions, concepts and methods to guide emerging research and practice*. GeoJournal, 72(3–4), 2008, 133–135. doi: 10.1007/s10708-008-9187-z
- Giddens, Anthony: *The constitution of society. Outline of the theory of structuration*. Oakland, CA: University of California Press, 1986
- Guo, Weisi; Donate, Guillem Mosquera; Law, Stephen; et al.: *Urban analytics: multiplexed and dynamic community networks*. arxiv:1706.05535v2 [cs.SI], 2017
- Kelley, Matthew James: *The emergent urban imaginaries of geosocial media*. GeoJournal, 78(1), 2013, 181–203. doi: 10.1007/s10708-011-9439-1
- Massey, Doreen: *Space, Place and Gender*. Cambridge, UK: Polity, 1994
- McFarlane, Colin: *The city as assemblage: dwelling and urban space*. Environment and Planning D: Society and Space, 29(4), 2011, 649–671. doi: 10.1068/d4710
- Meegan, Richard and Mitchell, Alison: *'it's not community round here, it's neighbourhood': neighbourhood change and cohesion in urban regeneration policies*. Urban Studies, 38(12), 2001, 2167–2194. doi: 10.1080/00420980120087117
- Mela, Alfredo; Chiodi, Sarah; Novascone, Roberta; et al.: *La città con-divisa. Lo spazio pubblico a Torino*. Milan, Italy: FrancoAngeli, 2014

Murdoch, Jon: *Post-structuralist geography: a guide to relational space*. London, UK: Sage, 2005. doi: 10.1111/j.1467-8306.2007.00587.x

Poorthuis, Ate and Zook, Matthew: *Artists and bankers and hipsters, oh my! Mapping tweets in the New York Metropolitan Region*. *Cityscape*, 16(2), 2014, 169–172

Reckwitz, Andreas: *Toward a theory of social practices: a development in culturalist theorizing*. *European Journal of Social Theory*, 5(2), 2002, 243–263

Rzeszewski, Michal: *Geosocial capta in geographical research—a critical analysis*. *Cartography and Geographic Information Science*, 45(1), 2018, 18–30. doi: 10.1080/15230406.2016.1229221

Schatzki, Theodore R: *Social practices. A Wittgensteinian approach to human activity and the social*. Cambridge, UK: Cambridge University Press, 1996. doi: 10.1017/CBO9780511527470

Semi, Giovanni: *Gentrification. Tutte le città come Disneyland*. Bologna, Italy: Il Mulino, 2015

Shelton, Taylor; Poorthuis, Ate; Graham, Mark; and Zook, Matthew: *Mapping the data shadows of hurricane Sandy: uncovering the sociospatial dimensions of 'big data'*. *Geoforum*, 52, 2014, 167–179. doi: 10.1016/j.geoforum.2014.01.006

Shelton, Taylor; Poorthuis, Ate; and Zook, Matthew: *Social media and the city: rethinking urban socio-spatial inequality using user-generated geographic information*. *Landscape and Urban Planning*, 142, 2015, 198–211. doi: 10.1016/j.landurbplan.2015.02.020

Shove, Elizabeth; Pantzar, Mika; and Watson, Matt: *The dynamics of social practice: everyday life and how it changes*. London, UK: Sage, 2012. doi: 10.4135/9781446250655

Soja, Edward W: *Postmodern geographies: the reassertion of space in critical social theory*. London, UK: Verso, 1989

Tuomela, Raimo: *The philosophy of social practices: a collective acceptance view*. Cambridge, UK: Cambridge University Press, 2002. doi: 10.1017/CBO9780511487446

Digital Imaginations of National Parks in Different Social Media: A Data Exploration

Vuokko Heikinheimo¹, Henrikki Tenkanen¹, Tuomo Hiippala^{1,2}, and
Tuuli Toivonen¹

¹*Digital Geography Lab, Department of Geosciences and Geography, University of Helsinki, Finland*

²*Department of Languages, University of Helsinki, Finland*

Social media contains a wealth of information about human activities in different places. This information can complement data collection efforts in resource-scarce fields such as nature conservation. However, social media platforms differ in popularity, content, and access to data, and the choice of platform may greatly affect the resulting analysis. We explored Flickr, Instagram, and Twitter data from 39 Finnish national parks over a period of two years to assess the fitness-for-purpose of each platform for understanding place-based experiences of national park visitors. From Instagram, we extracted data using two different approaches: coordinate search and keyword search. Furthermore, we identified the languages used in Instagram data using the fastText library, and conducted preliminary content analysis of Flickr and Twitter data using Google Cloud Vision image annotation service. Instagram was the most popular platform in all national parks. Noteworthy, almost 50% of Twitter users had shared their geotagged national park post to Twitter via Instagram. Language identification from text content and content analysis of images provide basis for further exploration of the digital representations of national parks and place-related experiences of visitors.

Keywords: social media; content analysis; national parks; Flickr; Instagram; Twitter

1 Introduction

Information available in Web 2.0 provides new opportunities for geographic knowledge discovery (Menis and Guo, 2009; Stefanidis et al., 2013; Sui and Goodchild, 2011) especially in data and resource-scarce fields such as nature conservation (Arts et al., 2015; Di Minin et al., 2015). Data on where, when, and why people visit and appreciate different places has become easier to gather in the era of big data, as people share their experiences and observations online in social media. Geosocial media platforms such as Flickr, Instagram, and Twitter contain a wealth of text, image, and video content about people's opinions, observations, activities, and experiences. However, social media platforms differ in popularity and purpose of use, and the choice of platform can greatly influence the observed patterns. Furthermore, spatial context including the cultural and physical environment plays a role in defining what content is shared and by whom.

Nature-based tourism has become increasingly popular around the world, with an estimated total of 8 billion visits per year to terrestrial protected areas (Balmford et al., 2015). Thus far, social media usage has been found to correlate with official visitor statistics of nature destinations (Hausmann et al., 2017; Tenkanen et al., 2017; Wood et al., 2013). Furthermore, it has been shown that geotagged social media content corresponds to surveyed activities (Heikinheimo et al., 2017) and preferences in terms of biodiversity (Hausmann et al., 2017) in selected national parks. Geosocial media has great potential

V Heikinheimo, H Tenkanen, T Hiippala, and T Toivonen (2018): *Digital Imaginations of National Parks in Different Social Media: A Data Exploration*. In: R Westerholt, F-B Moczniak, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 45–52

<https://doi.org/10.5281/zenodo.1472745>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

to inform protected area visitor monitoring and management (Di Minin et al., 2015), but more research is needed regarding the inherent bias in social media data and the suitability of different platforms for extracting information about place-related experiences in nature destinations.

Meanings associated with specific places have been extracted and analysed from user-generated content especially in urban environments with Twitter as the main data source (Jenkins et al., 2016; Shelton et al., 2015; Steiger et al., 2015). In environmental sciences, social media – especially Flickr and Panoramio – have been used for understanding landscape values and ecosystem services (benefits people get from nature) (Oteros-Rozas et al., 2016; Richards and Friess, 2015; Van Berkel et al., 2018; van Zanten et al., 2016). Recent studies have compared the spatial and temporal patterns of different social media to official statistics (Levin et al., 2017; Tenkanen et al., 2017). However, most studies – especially those focusing on content analysis – have so far relied on single source of social media (mostly Flickr or Panoramio in environmental studies, and Twitter in urban studies).

Social media content shared from natural and semi-natural areas has specific qualities in comparison with data from urban areas. Firstly, photos shared from protected areas are likely to be related to nature or cultural heritage. Secondly, most people who share content from national parks are likely to be visitors (domestic or international) enjoying their leisure time in the park (i. e., not living or working in the area). Also, some people might “log off” from social media during their visit or reduce their social media activities in order to save battery of their mobile device. People might use social media differently while being in a national park compared with built-up areas, and this should be taken into account when analysing these data.

Social media content can be attached to a place – in our case the national park – in different ways and at different scales. Coordinates and georeferenced place-tags (or points of interest) are the most technical way for the user to place their content on a map, and the precision and accuracy of geotagged data varies between platforms (Hochmair et al., 2018). It is also important to acknowledge that only a small percentage of all social media content is geotagged (it is estimated that 1% of all tweets are geotagged). Much of the place-related information is shared using place names and hashtags within text in a social media post, as is common in regular human discourse (Goodchild, 2011). Furthermore, images shared on social media also contain a wealth of place-related information.

In this short paper, we investigate how places of nature recreation – national parks in specific – are represented in different social media and discuss the possibilities and limitations of characterizing national parks (or people’s imaginations of national parks) from digital content. We compare and contrast data collected from Flickr, Instagram, and Twitter in terms of data volume, user base, and content. Furthermore, we analyse in more detail the text content of Instagram data in order to find out what language people use when sharing their experiences on social media. We also conducted preliminary analysis of image content from Flickr and Twitter. This data exploration provides a basis for more in-depth analysis of place-based experiences in recreational areas, and material for discussing the following questions: Which platforms are most suitable sources of place-related information from recreation areas? What is the best way of acquiring such data? Who have generated the digital data about a place?

2 Material and Methods

2.1 Study Area

This study covers all 39 Finnish national parks that existed in 2015. According to the international definition, national parks are large protected areas designed to protect ecosystems and to provide recreational opportunities to visitors (IUCN). In Finland, national parks are free of charge to everyone. Facilities such as campfire sites, nature trails, wilderness cabins, and latrines are maintained by the state-owned national park organization. In 2015, the 39 national parks attracted 2,634,600 visitors with a 15% increase from 2014 (<http://www.metsa.fi/kansallispuistotyhteensa>).

2.2 Data Collection

Spatial Search. We collected geotagged social media data from three different geosocial media platforms: Flickr, Instagram, and Twitter. Data was retrieved via the Application Programming

Table 1: Social media data from the 39 National Parks in 2014 and 2015.

Dataset	Posts	Users		
		total	min	max
Flickr	2 283	118	0	57
Instagram	7 627	4 137	4	1 308
Instagram, keyword search	110 176	42 931		
<i>geotagged</i>	<i>17 060</i>	<i>8 402</i>		
<i>geotagged in Finland</i>	<i>13 040</i>	<i>6 113</i>		
<i>geotagged in national parks</i>	<i>3 119</i>	<i>1 908</i>		
Twitter	5 567	729		
<i>excluding bots</i>	<i>3 979</i>	<i>728</i>	1	130
<i>tweets with any link</i>	<i>5 567</i>	<i>729</i>		
<i>source: www.instagram.com</i>	<i>653</i>	<i>356</i>		
<i>source: www.twitter.com</i>	<i>553</i>	<i>185</i>		
<i>thereof images</i>	<i>435</i>	<i>41</i>		
<i>source: www.swarmapp.com</i>	<i>100</i>	<i>64</i>		
<i>source: www.youtube.com</i>	<i>166</i>	<i>12</i>		

Interface (API) of each platform using a spatial query based on bounding boxes (Flickr and Twitter) or buffer zones around point locations (Instagram). For more details on the data collection, see methods presented by Tenkanen et al. (2017).

Keyword Search. From Instagram, another dataset was collected using a keyword-based media search with a manually collected list of place-names. Altogether, 504 place names at different spatial scales (ranging from trail names to park names in different languages) were included in the keyword list. All place names mentioned in visitor surveys and websites (<http://luontoon.fi>) of each national park were included in the list. The list was complemented with a manual search of place-names used in Instagram.

Pre-processing. All spatial datasets (Flickr, Instagram, and Twitter) were originally collected for larger areas, after which points intersecting the national park areas were selected. All datasets were subset to years 2014 and 2015, which was the most recent fully overlapping time period in all the collected datasets. Most active users (based on the number of posts per user) in each data set were manually checked to see if the data is generated by a machine, i. e., a bot. Posts generated by bots were removed.

2.3 Content Analysis

Source of Twitter Data. Twitter content was separately inspected in terms of data source. Any link shared on Twitter is automatically shortened by Twitter (<http://t.co>*) in order to reduce the number of characters in a tweet. This happens if users share a link in their tweet or have shared their post first in Instagram or other social networking sites (link to the original post will be available as a shortened link). The source of each shortened link was detected using a custom script (Python 3.5) and labelled according to their original source.

Languages. For Instagram data, we applied automatic language identification to each post using a pre-trained model via the fastText library (Bojanowski et al., 2017), which supports a total of 176 languages out-of-the-box. To pre-process the data, we followed the procedure set out by Hiiippala et al. (2018) to remove typical multilingual elements such as hashtags before retrieving predictions for each sentence in the caption. We then excluded very short sentences (< 7 characters) and predictions with a low confidence (< 0.5) from the results.

Image Content. Flickr and Twitter photo contents were labeled with up to 10 keywords using Google Cloud Vision (<https://cloud.google.com/vision/>) and the image annotation algorithm following the example of Richards and Tunçer (2017). Label detection was implemented in Python programming

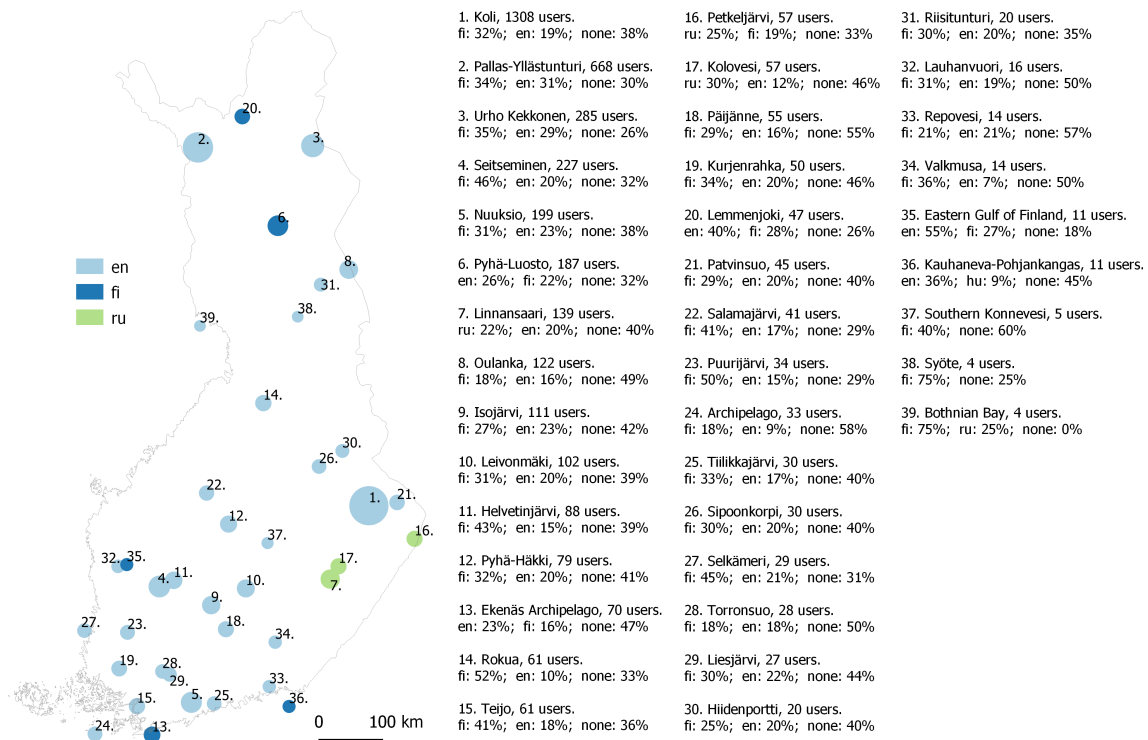


Figure 1: Map of the most popular languages among Instagram users in Finnish national parks. The proportion of the first and second most popular languages among users, as well as the proportion of users for whom language could not be detected (“none”) are presented separately for each national park. The abbreviations used for different languages are “fi” for Finnish, “en” for English, and “ru” for Russian. Symbol size represents the number of Instagram users within each park.

language using the Google Cloud Vision library. Photographs uploaded on Instagram were not automatically analysed due to restrictions introduced in 2018. Further analysis will be done to summarize these results at a later stage, e. g., hierarchical clustering following Richards and Tunçer (2017) and Oteros-Rozas et al. (2016).

3 Results

3.1 Data Volume

Coordinate Search. Instagram and Twitter contained information from each of the 39 national parks, Instagram being the most popular platform (Table 1). Flickr had the highest ratio of posts per user. Based on the manual check of most active users, Twitter was the only data source where automatically generated data, i. e., a bot, could be identified. The number of Instagram users for each national park are presented in Figure 1.

Keyword Search. The keyword search of Instagram data resulted in a relatively large dataset out of which 16% contained coordinate information. Out of these geotagged place-name search results, the majority (76%) were located within Finland, and almost one fifth (18%) had their coordinates within one of the national parks. The keyword search dataset and coordinate search dataset from Instagram had 1867 records in common (25% of coordinate-search data within national parks were found also via the keyword search).

Origins of Twitter Data. After excluding the most evident bot (a Twitter account posting automatically Finnish numbers with a random geotag), the Twitter dataset contained 3979 tweets by 728 users. Over 80% of the remaining tweets contained one or several URL-links. There were 210 different source websites, most common being instagram.com (content from Instagram), twitter.com (mostly photos

shared via Twitter), swarmapp.com (Foursquare check-ins) and youtube.com (videos). Over half of the Twitter users in this dataset had shared content originally generated in other social media platforms.

3.2 Content

Languages. Language could be identified for 56% of all captions (the rest were excluded due to length or low confidence associated with automatic language identification) and for 65% of all users in the coordinate-search-based Instagram data. Most captions were monolingual, i. e., they were written in a single language. Only 1% of the posts were bilingual. Across the entire dataset, Finnish and English were the dominant languages, followed by Russian and Swedish. Altogether, 31% of the captions were primarily or entirely written in Finnish, 21% in English, 6% in Russian, and 1% in Swedish. Popular languages across all national parks are visualized in Figure 1.

4 Discussion and Conclusion

In this paper we explored the volume and properties of social media data shared from Finnish national parks within a two year time period. Previous research has examined the relationship between the temporal patterns of social media activity and official visitor statistics (Tenkanen et al., 2017) across several parks, and compared the results of manual content classification and traditional surveys about national park visitor preferences and activities (Hausmann et al., 2017; Heikinheimo et al., 2017). This data exploration sets grounds for upcoming work where we will utilise automated content analysis methods to understand place-related experiences in national parks, and compare the results with official visitor information.

In terms of data volume, all of the platforms contained information from the observed national parks, but Instagram was clearly the most prominent source of social media data during the observed time period. Furthermore, Instagram usage correlates well with temporal visitation rates in the parks (Tenkanen et al., 2017). However, access to Instagram data through the Instagram API has been hindered since 2016, which affects the use of these data for scientific use. In addition to data availability, spatial context plays a role in determining which social media platform is most fit for capturing place-based experiences of people. While Flickr has been popular data source in environmental studies (Levin et al., 2017; Richards and Friess, 2015; Richards and Tunçer, 2017), it has clearly the smallest user-base among the studied platforms.

Interestingly, many of the geotagged tweets in our data set had originally been generated in other location-based social media platforms (Instagram, Foursquare, or Youtube). Furthermore, 10% of all tweets contained an image (shared originally in Twitter). Twitter is most often used as a source of text-based analysis (Jenkins et al., 2016; Steiger et al., 2015), but also the image content shared in Twitter (or in Instagram via Twitter) contains a wealth of useful information. This also means that some of the geotagged Instagram data can still be accessed programmatically, despite the changes in the Instagram API.

We used two different approaches for collecting Instagram data from the study areas: a coordinate-based spatial query and a keyword-based place-name query. Unsurprisingly, the keyword search captured a lot of data that was not related to our areas of interest due to keywords with multiple purposes. For example, the keyword “Koli” is the name of a popular national park in Eastern Finland but also a language dialect in India, among other meanings. Trough the combination of the place-name query and a spatial query (selecting keyword search results within Finland or within National Park borders), we probably managed to exclude a lot of irrelevant content located outside Finnish national parks, but also potentially lost relevant non-geotagged information and relevant content posted outside the borders of the national parks. As Goodchild (2011) has argued, an intelligent strategy would be needed in order to develop a search which captures meanings of different places and place names correctly from human-generated content.

Automatic language identification helps us to understand who is sharing content in social media and for which audience. As such, geotagged language information may reveal the linguistic landscape of an area and give hints about the origins of visitors (Hiippala et al., 2018). In our case, the simple language detection revealed that national parks as largely visited by national visitors. Language identification is also a crucial pre-processing step for further text analysis methods, such as topic

modelling and sentiment analysis. In conjunction with other types of user-related information, language helps us to better understand who has generated the data and for what purpose, and whose digital imaginations of a place are represented in social media.

Social media data analysis from well-monitored national parks has the potential to provide new information about gaps and bias in different social media data (Hausmann et al., 2017; Tenkanen et al., 2017; Wood et al., 2013). National park visitor surveys and visitor counting serve as ground-truth information for patterns observed in online social media. Our further work will focus on comparing online social media content to experiences measured with more traditional methods such as visitor surveys across multiple parks.

Using location-based social media data in research requires constant reflection about data quality and research ethics (boyd and Crawford, 2012). Here (and in many other studies) social media has been mined from online sources without specific consent from users who have generated the data. However, just the fact that these data are online does not necessarily justify their use in respect to a new purpose (boyd and Crawford, 2012). In research, one should constantly consider the potential benefit and harm to anyone involved, and to ensure the protection of personal information (Monkman et al., 2017).


In sum, understanding place-based experiences from social media can benefit from applying existing machine learning methods as well as from the development of new, more efficient ways of automatically extracting and analysing place-related information. Well-monitored national parks, such as those in Finland, provide a convenient test environment, and an interesting application case for observing collective experiences and emerging phenomena from geosocial media. In future work we aim to deepen general understanding about place-related experiences of people in national parks based on social media data, and to provide further insights about which platforms are most suitable for extracting these information, what is the best way of acquiring such data, and who have generated the digital data about a place.


Funding


We thank the Kone Foundation for support.

ORCID

Vuokko Heikinheimo  <https://orcid.org/0000-0001-5119-0957>

Henrikki Tenkanen  <https://orcid.org/0000-0002-0918-4710>

Tuomo Hiippala  <https://orcid.org/0000-0002-8504-9422>

Tuuli Toivonen  <https://orcid.org/0000-0002-6625-4922>

References

Arts, Koen; van der Wal, René; and Adams, William M: *Digital technology and the conservation of nature*. *Ambio*, 44, 2015, 661–673. doi: 10.1007/s13280-015-0705-1

Balmford, Andrew; Green, Jonathan MH; Anderson, Michael; et al.: *Walk on the wild side: estimating the global magnitude of visits to protected areas*. *PLoS Biology*, 13(2), 2015, e1002074. doi: 10.1371/journal.pbio.1002074

Bojanowski, Piotr; Grave, Edouard; Joulin, Armand; and Mikolov, Tomas: *Enriching word vectors with subword information*. *Transactions of the Association for Computational Linguistics*, 5, 2017, 135–146

boyd, danah and Crawford, Kate: *Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon*. *Information, Communication & Society*, 15(5), 2012, 662–679. doi: 10.1080/1369118X.2012.678878

Di Minin, Enrico; Tenkanen, Henrikki; and Toivonen, Tuuli: *Prospects and challenges for social media data in conservation science*. *Frontiers in Environmental Science*, 3, 2015. doi: 10.3389/fenvs.2015.00063

Goodchild, Michael F: *Formalizing place in geographic information systems*. In: Burton, Linda M; Matthews, Stephen A; Leung, ManChui; Kemp, Susan P; and Takeuchi, David T (eds.), *Communities, neighborhoods, and health: expanding the boundaries of place*, New York, NY: Springer, 2011. 21–34. doi: 10.1007/978-1-4419-7482-2_2

Hausmann, Anna; Toivonen, Tuuli; Slotow, Rob; et al.: *Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas*. *Conservation Letters*, 11(1), 2017. doi: 10.1111/conl.12343

Heikinheimo, Vuokko; Minin, Enrico Di; Tenkanen, Henrikki; et al.: *User-generated geographic information for visitor monitoring in a national park: a comparison of social media data and visitor survey*. *ISPRS International Journal of Geo-Information*, 6(3), 2017, 85. doi: 10.3390/ijgi6030085

Hiippala, Tuomo; Hausmann, Anna; Tenkanen, Henrikki; and Toivonen, Tuuli: *Exploring the linguistic landscape of geotagged social media content in urban environments*. *Digital Scholarship in the Humanities*, 2018. doi: 10.1093/lc/fqy049

Hochmair, Hartwig H; Juhász, Levente; and Cvetojevic, Sreten: *Data quality of points of interest in selected mapping and social media platforms*. *Proceedings of the 14th International Conference on Location Based Services*, 2018, 293–313. doi: 10.1007/978-3-319-71470-7_15

Jenkins, Andrew; Croitoru, Arie; Crooks, Andrew T; and Stefanidis, Anthony: *Crowdsourcing a collective sense of place*. *PLoS ONE*, 11(4), 2016, e0152932. doi: 10.1371/journal.pone.0152932

Levin, Noam; Lechner, Alex Mark; and Brown, Greg: *An evaluation of crowdsourced information for assessing the visitation and perceived importance of protected areas*. *Applied Geography*, 79, 2017, 115–126. doi: 10.1016/j.apgeog.2016.12.009

Mennis, Jeremy and Guo, Diansheng: *Spatial data mining and geographic knowledge discovery – an introduction*. *Computers, Environment and Urban Systems*, 33(6), 2009, 403–408. doi: 10.1016/j.compenvurbsys.2009.11.001

Monkman, Graham George; Kaiser, Michel; and Hyder, Kieran: *The ethics of using social media in fisheries research*. *Reviews in Fisheries Science & Aquaculture*, 26(2), 2017, 235–242. doi: 10.1080/23308249.2017.1389854

Oteros-Rozas, Elisa; Martín-López, Berta; Fagerholm, Nora; Bieling, Claudia; and Plieninger, Tobias: *Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites*. *Ecological Indicators*, 2016. doi: 10.1016/j.ecolind.2017.02.009

Richards, Daniel R and Friess, Daniel A: *A rapid indicator of cultural ecosystem service usage at a fine spatial scale: content analysis of social media photographs*. *Ecological Indicators*, 53, 2015, 187–195. doi: 10.1016/j.ecolind.2015.01.034

Richards, Daniel R and Tunçer, Bige: *Using image recognition to automate assessment of cultural ecosystem services from social media photographs*. *Ecosystem Services*, 31(C), 2017, 318–325. doi: 10.1016/j.ecoser.2017.09.004

Shelton, Taylor; Poorthuis, Ate; and Zook, Matthew: *Social media and the city: rethinking urban socio-spatial inequality using user-generated geographic information*. *Landscape and Urban Planning*, 142, 2015, 198–211. doi: 10.1016/j.landurbplan.2015.02.020

Stefanidis, Anthony; Crooks, Andrew; and Radzikowski, Jacek: *Harvesting ambient geospatial information from social media feeds*. *GeoJournal*, 78(2), 2013, 319–338. doi: 10.1007/s10708-011-9438-2

Steiger, Enrico; Westerholt, René; Resch, Bernd; and Zipf, Alexander: *Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data*. *Computers, Environment and Urban Systems*, 54, 2015, 255–265. doi: 10.1016/j.compenvurbsys.2015.09.007

Sui, Daniel and Goodchild, Michael F: *The convergence of GIS and social media: challenges for GI-Science*. *International Journal of Geographical Information Science*, 25(11), 2011, 1737–1748. doi: 10.1080/13658816.2011.604636

Tenkanen, Henrikki; Di Minin, Enrico; Heikinheimo, Vuokko; et al.: *Instagram, Flickr, or Twitter: assessing the usability of social media data for visitor monitoring in protected areas*. Scientific Reports, 7(17615), 2017. doi: 10.1038/s41598-017-18007-4

Van Berkel, Derek B; Tabrizian, Payam; Dorning, Monica A; et al.: *Quantifying the visual-sensory landscape qualities that contribute to cultural ecosystem services using social media and LiDAR*. Ecosystem Services, 31, 2018, 326–335. doi: 10.1016/j.ecoser.2018.03.022

Wood, Spencer A; Guerry, Anne D; Silver, Jessica M; and Lacayo, Martin: *Using social media to quantify nature-based tourism and recreation*. Scientific Reports, 3, 2013, 2976. doi: 10.1038/srep02976

van Zanten, Boris T; Van Berkel, Derek B; Meentemeyer, Ross K; et al.: *Continental-scale quantification of landscape values using social media data*. Proceedings of the National Academy of Sciences, 113(46), 2016, 12974–12979. doi: 10.1073/pnas.1614158113

From Space to Place and Back Again: Towards an Interface Between Space and Place

Emmanuel Papadakis¹, George Baryannis², and Thomas Blaschke¹

¹Department of Geoinformatics – Z_GIS, University of Salzburg, Austria

²Department of Computer Science, University of Huddersfield, UK

Geographic information systems represent and process space whereas people refer to and use place. A question that arises is what are the benefits of introducing a unified data model that combines the rigid representation of space and the information-rich concepts of place. In this work we contribute to this research question by proposing a two-way interface that aims to bridge the notions of space and place. This interface relies on the four conceptions of space and interconnected spatial objects. Step-by-step descriptions as well as examples are provided to illustrate the intended use of the proposed interface.

Keywords: space; place; place-based GIS; objectification of space; ontology

1 Introduction

Geographic information science in combination with natural and social sciences utilize the notions of space and place to describe the geographical world. Space and place see the cosmos through two perspectives: a) as depicted by the rigorous methods of mathematics and physics, and b) as interpreted by the complexity of human mentality, respectively. Representation attempts for space and place have resulted in approaches from both sides that offer adequate descriptions of the world. On the one hand, “puzzles of polygons and sandwiches of data layers” (Couclelis, 1992b) introduce a spatial view that adheres to objectivity, concise abstractions, and numeric precision. On the other hand, complicated models, semantic enrichment, and stochastic approaches emphasize on a place-based illustration that allows subjectivity, extensive expressiveness, and increased vagueness (Goodchild, 2011). An emerging question is whether these disparate views of the world should be kept separated and which are the benefits and impacts of unifying them. In other words, shall the space and place representation methods keep their individuality and “excel” on providing particular descriptions of the world or should they be combined introducing a unified view of the world that includes both spatial and “platial” features?

People live and act in the geographic world. Their intrinsic need to describe and talk about it led to the creation of symbols, i. e., space and place. The connotation of those terms, which allows the scientific foundation around them, originates from the vernacular. Considering that languages differ based on the culture of the people who use them, the problem of converging to a widely acceptable interpretation is not to be neglected. However, the influence of the ancient Greek philosophy played an important role on sculpting the meaning of space and place in the modern Western world’s mentality (Purves, 2010). Plato and Aristotle introduced the concept of space as a geometrical notion disconnected from matter (Beichler, 1981). Analysing the ancient Greek narratives, Algra (1997) suggests that space has three possible interpretations

- a geometrical extent occupied by an entity,

E Papadakis, G Baryannis, and T Blaschke (2018): *From Space to Place and Back Again: Towards an Interface Between Space and Place*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 53–59

<https://doi.org/10.5281/zenodo.1472747>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

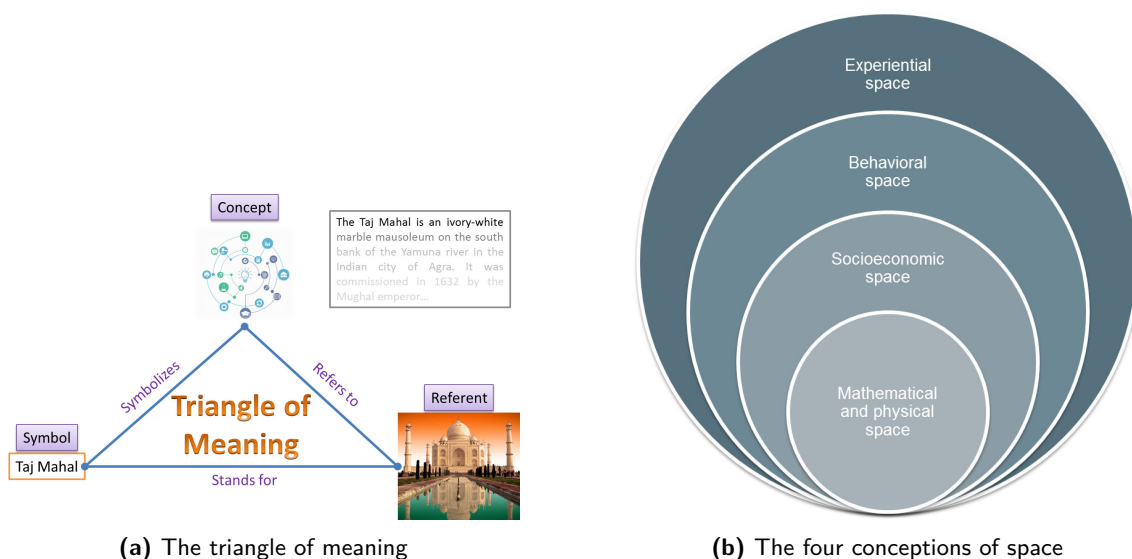


Figure 1: Meaning and space

- a location of an entity, and
- a container that includes or excludes entities.

On the contrary, place most often combines relative spatial information and association with other entities. Algra (1997) argues that place is perceived as

- a relative extent of an entity,
- a relative location of an entity, and
- relevant information that go beyond the entity's extent and location.

Both the notions of space and place describe the world in relation with the a) extension, b) localization, and c) containment of entities. However, the way those entities are perceived by an observer suggests different perspectives. On the one hand, space refers to the geometrical properties of entities, e. g., “a box occupies $10 \times 17 \times 20 \text{ cm}^3$ ”, “a box is located at (0,0)”, and “the entities that exist within the box's volume”. On the other hand, place focuses on the concepts and associations that build the entities of the world. For example, “the box occupies the whole room”, “the box is in the centre of the room”, and “laptop boxes are stored in the electronic devices room”. Through the spatial perspective a “box” is considered as a geometrical shape and its spatial information is quantitative. Place, however, focuses on the thematic information and relations that hold for the entities under consideration. In this case, the “box” and the “room” are whole and particular conceptual entities carrying their own meaning; localization and extension emphasize on their relative spatial association, whereas containment implies the semantic association that a laptop is an electronic device.

Geographic information systems (GIS) represent space using the basic standard data structures, vector and raster (Frank, 1992b), setting it in relation with a coordinate reference system. Space conforms to a strict formalization that allows quantification and implies objectivity and precision. These features make space transferable and processable by machines. On the other hand, people refer to place as a consequence of their qualitative perception of the world (Frank, 1992a). Place is a human invention to describe the geographical world (Curry, 1996) and is built by combining experience with spatial information (Couclelis, 1992a). In other words, place is space infused with human meaning (Tuan, 1979). Based on the aforementioned definition and considering the triangle of meaning (Figure 1a) a place is modelled with concepts that describe meaningful entities of the real world extended with relative spatial information, which takes the form of relations and semantic associations. Since place is a product of human thinking, it is subjective, vague, and informal.

Considering the above, machines represent and process space and associated representations, whereas people tend to use and refer to place. Particularly, in spite of space being used in the human discourse, GIS representation standards make sense only within fields of natural and social sciences.

On the other hand, models that describe place are mainly products of philosophical discussion or theoretical models that lack formalisation. Therefore, such models are not machine-interpretable, which hinders their ability to assist in spatial processing.

In order to address the above difficulties, we propose an interface between GIS spatial representation standards and theoretical platial models. This interface is founded upon two influential strands of GIS research: First, the interconnection between space and place, as analysed by Relph (1976), and secondly, the four conceptions of space, as proposed by Couclelis (1992a). We envision that this proposal can contribute to the ongoing efforts of the research community to bridge space and place together.

2 Background and Related Work

We first summarize the works that are the foundation of the proposed interface. Relph (1976) construes place as a unique pattern of physical features, appearances, activities, and functions. Its unique quality is the power to focus on human intentions, experiences, and actions in the spatial dimension. Relph (1976) also mentions the close association of place and space, stating that place is built based on environmental experience, which, in turn, is derived by the the spatial context of the place.

Couclelis (1992a) introduced a multilevel composition of place to reveal the inclusion of space into place, in the form of four conceptions illustrated in Figure 1b. The core concept refers to the mathematical space as defined by mathematical and physical laws; it is the space that is used in quantitative geography, including absolute and relativistic space. The next level introduces the notion of relative space, denoted as socioeconomic space. In this concept, characteristic properties that yield social or economical interests are analysed and represented spatially. Moving to a lower level of formalism comes the concept of behavioural space, which is determined by spatial cognition and understanding. It is an incomplete, distorted, and highly subjective space that depicts the spatial decisions and the behaviour based on the individual's knowledge and psychology. The last concept describes the experiential space. It refers to the perceived world, prior to filtering and scientific analysis.

In the rest of this section, we briefly summarize related efforts to bridge the gap between the notions of space and place. The field of quality of life introduces techniques that convert the nominal qualitative values of abstract human data to system-friendly representations, such as numbers (Nussbaum and Sen, 1993), by using statistics and GIS methods. This is made possible through the qualitative space approach (Frank, 1992a). This method is considered as semantic enrichment that focuses on the assignment of qualitative properties on spatial regions. Ontological gazetteers (Hill, 2000) take the idea of semantic enrichment one step further by offering spatially-referenced catalogues of place names with the ability of ontological orientation. They provide a linkage between the human and physical world by encoding relations between place names, space footprints, spatial categories, and so on.

The semantic enrichment attempted by both quality of life and ontological gazetteers research does not fully represent the expressiveness of human meaning, lacking relational semantics and any information that make up the object-based view of place. The latter view is the focus of the more recent line of research on objectification of space (Couclelis, 1992b; Goodchild et al., 2007). Particularly, people tend to simplify complicated spatial structures into compact sophisticated objects with certain properties and functions. In other words, people think with objects. For instance, a mountain is represented as a field in a GIS revealing its elevation model and so on. However, people consider it as an object, with certain properties such as name, location, and overall height.

Fully-fledged place ontologies have also been developed in the context of linked data, such as GeoNames¹ or Schema.org². Furthermore, the concept of semantic place (Scheider and Purves, 2013) augments ontological formalization with relational semantics extracted from narratives, to determine relations among explicit places, implicit places, or other entities. However, since their focus is place localization, semantic places lack geometric information, instead compensating with relative spatial relations and extent.

A data-driven approach to moving from space to place is conducted by Gao et al. (2017b), with the objective to approximate the spatial extent and thematic information of a cognitive region (Hollenstein and Purves, 2010) using probabilistic models derived from topic modelling algorithms. A relevant data-driven approach is presented by Gao et al. (2017a), with the authors employing probabilistic models to determine the semantic signature of functional regions. These approaches introduce formalization methods that deal with the vagueness of place definition; however, their reliance on data often leads to

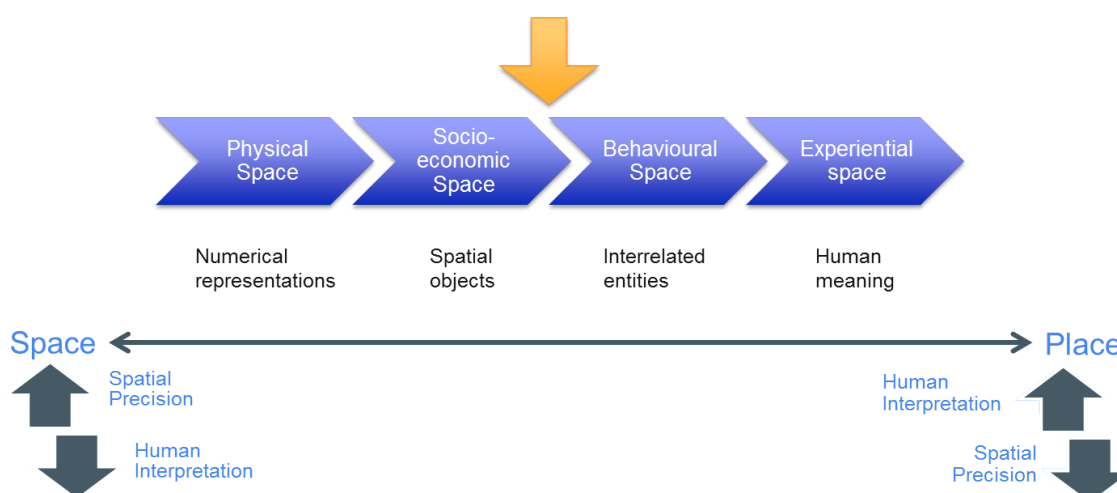


Figure 2: The golden mean between space and place

results that may not be comprehensible or sufficiently relevant to the human meaning that is attached to places.

3 An Interface Between Space and Place

The proposed interface is based on the inter-dependency of space and place (Relph, 1976) to form a bridge between the two notions. In both cases, suitable abstractions allows place and space to acquire some essential characteristics of the respective other by giving up some of their own features, such as precision in the case of space and expressiveness in the case of place. To achieve this, we use the conceptions of space by Couclelis (1992a). As illustrated in Figure 2, space maps directly to the leftmost conception (physical space), while place is the rightmost one (experiential space). The ultimate goal of the interface is to move from these two extremes towards the two intermediate conceptions, socioeconomic and behavioural space. A combination of both would be akin to a golden mean between space and place: objects and relations are used to represent space in a form that is understandable by humans but, at the same time, these objects are spatially projected, providing the precision required by GIS. However, this depends on the nature of each application and the availability of suitable data.

The interface is expected to work in two directions, from place to space and from space to place, the latter of which is relatively less complicated and is described first. Working in the direction from space to place, the interface is given a spatial representation as an input and is first tasked to infuse it with aspects of meaning converting it to a human-interpretable format (Tuan, 1979). This, for instance, can be in the form of fiat or bona-fide spatial objects (Cova and Goodchild, 2002), each of which is associated with a descriptive symbol, as well as relevant thematic information. This would correspond to socioeconomic space. The interface then depends on the availability of defined relations among spatial objects in order to move towards behavioural space.

Working in the other direction, given an instance of a model of place, the interface should decompose the place into aspects related to location, affordances, and equipment. Then, the interface should abstract away purely subjective information from these aspects, such as emotions or activities. This can be achieved by clustering such features into categories in order to facilitate formalization. Afterwards, any information that is in an unstructured form (e. g., narratives) has to be adapted into a structured form, such as a knowledge graph. At this point, we reach behavioural space through the formalized representation of the place's equipment. Finally, location and extent information attached to the place at hand are employed to convert the place's equipment to spatial objects through the introduction of an appropriate system of coordinates. This would require a form of fuzzification, since places tend to have fiat boundaries. This step brings us to socioeconomic space, concluding the work of the interface.

It should be noted that the aforementioned analysis is maximal, including all possible steps in both directions. However, depending on the space or place model given as input, some of these steps

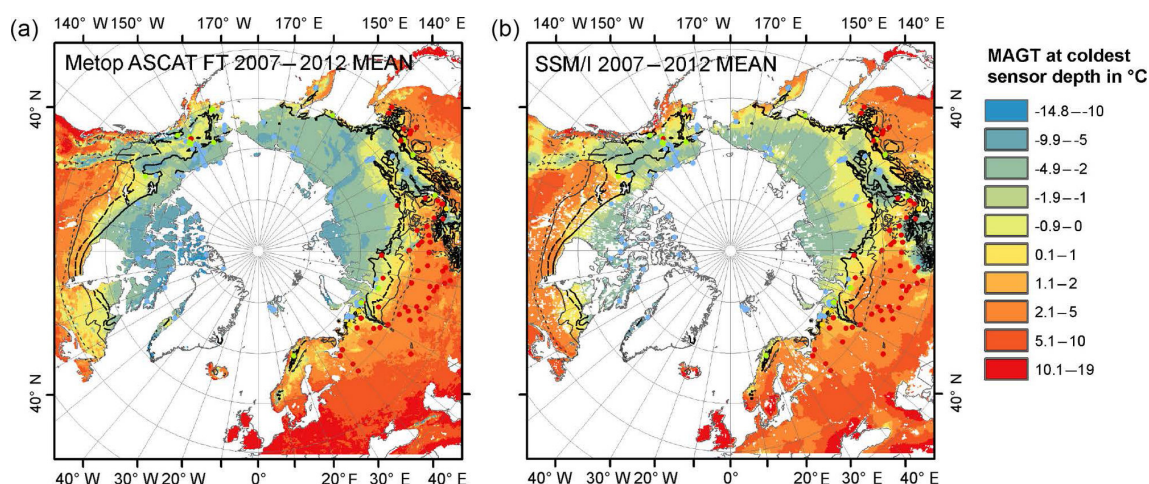


Figure 3: Maps of modelled mean annual ground temperature (Kroisleitner et al., 2018); use permitted under the Creative Commons Attribution License CC BY 4.0 (<http://creativecommons.org/licenses/by/4.0/>)

may not be necessary or can be simplified. For instance, if the model of place provided to the interface is in the form of an ontology, then it is already represented in a structured form. Also, if the input to the interface is qualitative space, then the conversion to spatial objects is straightforward.

4 Examples

In this section, we present brief examples of applying the proposed interface in both directions. Figure 3 shows maps of modelled annual ground temperature used in permafrost detection, an example of purely spatial information extracted by satellite imagery and semantically enriched by sensor data. Such information is represented by aggregated points plotted over a map and does not fully conform to the human perception of place. Using the proposed interface, the continuous temperature spectrum is converted to a set of categorical values, namely permafrost extent classes. Then, areas associated with a common categorical value introduce spatial objects with fuzzified boundaries, which correspond to socioeconomic space. Any additional information available for these areas that enable spatial cognition, such as place names from the vernacular, allows us to move towards behavioural space; for instance, “parts of Ellesmere island have continuous permafrost extent”. Introducing spatial relations for the aforementioned spatial objects allows the establishment of associations among them, e. g., “the area south of Mount Whisler has a scarce permafrost extent”. The features provided by the behavioural space allow spatial reference and localization using comprehensible terms instead of arbitrarily regions defined by clustering methods.

An example moving from place to space follows. Given a model of an airport place in the form of a narrative³ as an input, the proposed interface decomposes it into location, affordances, and equipment aspects, as shown in Figure 4a. Afterwards, these features are clustered into categories, e. g., the narrative may provide equipment information about several different facilities such as terminals, towers, and hangars, which are clustered into a category of buildings. All extracted categories along with relations among them are then formalized into a hierarchical ontology (Figure 4b), reaching behavioural space. Finally, combining the semantic relations within the ontology along with additional information about design standards or regulations, a spatial projection on a geometric space is introduced. If any additional location information is available (e. g., the absolute location of an airport in London), the spatial projection of the airport as an object can be mapped to the geographic space. Either a geometric or a geographic space projection corresponds to socioeconomic space.

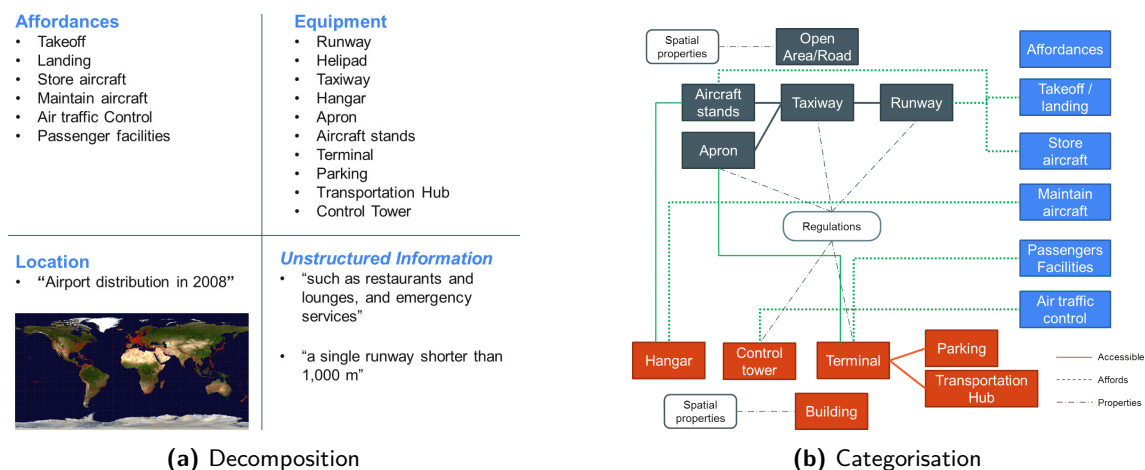


Figure 4: Using the interface on an airport place

5 Conclusion

In this paper, we have proposed to bridge space and place through a structured two-way interface based on the four conceptions of space by Couclelis (1992a). This interface enables a form of representation, based on interconnected spatial objects. Such a representation is understandable by humans but, at the same time, includes spatial projection, providing the precision required by GIS. This proposal aims to stimulate researchers working on the confluence of space and place to move one step closer to data models that achieve both of the following: First, to allow place to be involved in spatial analysis and processing and, secondly, to make complex spatial representations more easily interpretable by humans.

Future steps include coming up with a formal definition of the interface that offers strict guidelines on the conversion between space and place data models. This formal definition should then be validated using complex real-world examples.

Notes


1. <http://www.geonames.org/ontology>
2. <https://schema.org/Place>
3. An example narrative is the Wikipedia definition of an aerodrome (<https://en.wikipedia.org/wiki/Aerodrome>)


Acknowledgements

The presented work is framed within the Doctoral College GIScience (DK W 1237N23), funded by the Austrian Science Fund (FWF).

ORCID

Emmanuel Papadakis  <https://orcid.org/0000-0001-8669-2420>

George Baryannis  <https://orcid.org/0000-0002-2118-5812>

Thomas Blaschke  <https://orcid.org/0000-0002-1860-8458>

References

- Algra, Keimpe: *Concepts of space in Greek thought*. Leiden: Brill, 1997
- Beichler, James E: *The ancient Greek concept of space*. https://www.academia.edu/8737012/The_Ancient_Greek_Concept_of_Space, 1981. Retrieved 15 September 2018

Couclelis, Helen: *Location, place, region, and space*. In: Abler, Ronald F; Marcus, Melvin G; and Olson, Judy M (eds.), *Geography's inner worlds*, New Brunswick, NJ: Rutgers University Press, 1992a. 215–233

— *People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS*. Proceedings of the International Conference GIS, From Space to Territory: Theories and Methods of Spatio-Temporal Reasoning, 1992b, 65–77. doi: 10.1007/3-540-55966-3_3

Cova, Thomas J and Goodchild, Michael F: *Extending geographical representation to include fields of spatial objects*. *International Journal of Geographical Information Science*, 16(6), 2002, 509–532. doi: 10.1080/13658810210137040

Curry, Michael R: *The work in the world: geographical practice and the written word*. Minneapolis, MN: University of Minnesota Press, 1996

Frank, Andrew U: *Qualitative spatial reasoning about distances and directions in geographic space*. *Journal of Visual Languages & Computing*, 3(4), 1992a, 343–371. doi: 10.1016/1045-926X(92)90007-9

— *Spatial concepts, geometric data models, and geometric data structures*. *Computers & Geosciences*, 18(4), 1992b, 409–417. doi: 10.1016/0098-3004(92)90070-8

Gao, Song; Janowicz, Krzysztof; and Couclelis, Helen: *Extracting urban functional regions from points of interest and human activities on location-based social networks*. *Transactions in GIS*, 21(3), 2017a, 446–467. doi: 10.1111/tgis.12289

Gao, Song; Janowicz, Krzysztof; Montello, Daniel R; et al.: *A data-synthesis-driven method for detecting and extracting vague cognitive regions*. *International Journal of Geographical Information Science*, 31(6), 2017b, 1245–1271. doi: 10.1080/13658816.2016.1273357

Goodchild, Michael F: *Formalizing place in geographic information systems*. In: Burton, Linda M; Matthews, Stephen A; Leung, ManChui; Kemp, Susan P; and Takeuchi, David T (eds.), *Communities, neighborhoods, and health*, New York, NY: Springer, 2011. 21–33. doi: 10.1007/978-1-4419-7482-2_2

Goodchild, Michael F; Yuan, May; and Cova, Thomas J: *Towards a general theory of geographic representation in GIS*. *International Journal of Geographical Information Science*, 21(3), 2007, 239–260. doi: 10.1080/13658810600965271

Hill, Linda L: *Core elements of digital gazetteers: placenames, categories, and footprints*. Proceedings of the International Conference on Theory and Practice of Digital Libraries, 2000, 280–290. doi: 10.1007/3-540-45268-0_26

Hollenstein, Livia and Purves, Ross: *Exploring place through user-generated content: using Flickr tags to describe city cores*. *Journal of Spatial Information Science*, 2010(1), 2010, 21–48. doi: 10.5311/JO-SIS.2010.1.3

Kroisleitner, C; Bartsch, A; and Bergstedt, H: *Circumpolar patterns of potential mean annual ground temperature based on surface state obtained from microwave satellite data*. *The Cryosphere*, 12(7), 2018, 2349–2370. doi: 10.5194/tc-12-2349-2018

Nussbaum, Martha and Sen, Amartya: *The quality of life*. Oxford, UK: Oxford University Press, 1993

Purves, Alex C: *Space and time in ancient greek narrative*. Cambridge, UK: Cambridge University Press, 2010

Relph, Edward: *Place and placelessness*. London, UK: Pion, 1976

Scheider, Simon and Purves, Ross: *Semantic place localization from narratives*. Proceedings of the 1st ACM SIGSPATIAL International Workshop on Computational Models of Place (COMP'2013), 2013. doi: 10.1145/2534848.2534858

Tuan, Yi-Fu: *Space and place: humanistic perspective*. In: Gale, Stephen and Olsson, Gunnar (eds.), *Philosophy in Geography*, Dordrecht, The Netherlands: Springer, 1979. 387–427. doi: 10.1007/978-94-009-9394-5_19

The Value of Detours

Sanam N Vardag¹ and Sven Lautenbach^{2,1}

¹Heidelberg Center for the Environment, Heidelberg University, Germany

²Institute of Geography, Heidelberg University, Germany

The estimation of the value of georeferenced spaces is challenging because the value of a space depends on many factors such as recreation potential, sociability, cultural points of interest, etc. To account for these personal values, interviews and surveys can be conducted but this is costly and elaborate, and this requires a user action. Recent approaches use social media to attribute positive or negative associations of users to a georeferenced space by analysing the semantic linkages. This approach is promising but at the same time prone to biases due to a user selection bias and the interpretation of the semantics. In this paper, an alternative method based on the systematic analysis of georeferenced paths of people is proposed, which does not require direct user interaction. Underlying the assumption that people are willing to take a longer and more time-consuming path if the detour has a personal value for them, this work proposes to estimate these personal values systematically and partly automated by analysing the pathways taken. The authors suggest to conduct a first feasibility study analysing the GPS positions of people to obtain information and find patterns of the personal preferences of places, and matching them to interview or social media derived information on platial preferences.

Keywords: cultural values; routing; detours; urban green spaces

1 Introduction

What is the value of spiritual Buddhist water? What is the value of the lake where you had your first kiss? What is the value of the dirty lake, which you heard might be contaminated? What is the value of going water-skiing every Monday? And what if all four descriptions refer to same lake? The value of a place depends on the associations that are connected to it and to the phase room of possibilities, which it offers to the users. The concept of platial goes beyond the actual space in a geographic sense, but couples the space with place names, descriptions, and semantic relationships between places (Gao et al., 2013). Elaborating on this concept, it is often desirable to even include the social and cultural possibility to interact with the space. The lake becomes more than pure coordinates and metrics – it becomes a concept of its own. How can one include the social and cultural dimension into the geoinformation? Recent work suggest using social media such as Flickr, Twitter, Facebook, Spotify, etc. for attributing positive or negative associations of users to a georeferenced space by analysing the semantic linkages (Coscieme, 2015; Gliozzo et al., 2016; Lee et al., 2018; Oteros-Rozas et al., 2016; Richards and Tunçer, 2018; Yoshimura and Hiura, 2017). For example, in Twitter this may work by analysing georeferenced hashtags, which are in the same post as #lake (or to be more specific #lakegeneva). Positive hashtags such as #wonderful #greatday #unforgettable signify a large personal value of the lake, whereas negative hashtags such as #dirty #deadfish or #seenbetter signify a negative personal value to the lake. This analysis couples positively and negatively associated words with the coordinates incorporating a social labelling. However, it also inhibits some uncertainties. One uncertainty may be due to a misinterpretation of the positiveness or negativeness of some hashtags. For example, a post with hashtags #dirty #Hands #cleaned #in #Lakegeneva #thanksbeautifulnature

SN Vardag and S Lautenbach (2018): *The Value of Detours*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18), pp. 61–65

<https://doi.org/10.5281/zenodo.1472749>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

would have the negative hashtag #dirty in it even though the post actually associates a positive value to the lake. A careful selection of posts and hashtags, also in combination with each other, is therefore needed for interpretation. Further, the people using social media are not representative of the entire community, as in particular young people use social media (Mellon and Prosser, 2017). Therefore, care must be taken when drawing conclusions about the whole underlying population. Depending on the application it might be an option to treat values derived from social media as presence-only data. Otherwise it is necessary to limit the conclusions only to the relevant peer group.

In this paper, a new method based on the systematic analysis of detours is presented, which seeks to include the social cultural dimension of a georeferenced space. A detour is defined as a “different or less direct route to a place that is used to avoid a problem or to visit somewhere or do something on the way” (Cambridge University Press, 2018). If one decides to take a detour, the perceived benefit of the integrated detour path must outweigh the perceived additional cost of a longer distance and typically a longer time effort. In other words, the length and time demand of the detour can be used to quantify the service provided by green spaces but also social places by means of a travel cost approach. Therefore, the choice of way contains information on the personal value which people attribute to the integrated possible paths one can take. As an example, when going from home to work, one can decide to take the fastest way. However, one can also decide to take a detour. Why would one do this? It might be that there is a bakery on the way or to simply enjoy taking the longer but more scenic or greener way, or that the shortest way is dangerous or ugly. All of these reasons assign a relative value to the integrated detour chosen relative to the integrated fastest actual path from A to B. Systematically analysing the detours chosen by a group of people may therefore provide insight into the ratio of different groups of people and may allow us to understand the personal values of the integrated paths.

2 Methods

Our aim is to estimate the personal value of urban green spaces and social places. The principle idea is to analyse, which detour a person is willing to take when going from A to B. This enables an estimation of the perceived benefits of the integrated path, which must compensate the perceived additional costs, i. e., additional time needed for the detour. By knowing the start and endpoint of the tour and therewith the fastest possible route as well as the actual taken pathway, one can learn about the personal value, which a person attributes to the chosen way that was chosen in comparison to the fastest possible. Since we have no information why a specific route was chosen we have to rely on the statistical analysis of a large number of different detours and their properties such as greenness of the route or number of places of interest along the route. To our knowledge the automated detour detection is novel as it does not require the user to recognize why and that he or she prefers a certain route as is the case in user surveys and social media. Therefore the proposed method may be an ideal complement for time-consuming and actively induced information on places.

We plan to test our proposed method in a first feasibility study. The analysis requires a high resolution of GPS data of about 1 minute to differentiate the mean of transportation and any stop in the mean time, a sufficiently high and representative sample size over the specified area, and a sufficiently long time duration of GPS tracking to rule out effects of weather conditions etc. Finally, in an ideal setting we are able to match and cross-check our findings with a bottom-up approach (i. e., user surveys and social media) conducted with the same participants. A possible data set for this analysis could be provided within the Psychogeography Project¹, in which GPS sensors as well as GPS triggered e-diaries are used to capture the position and mood of the participants by self-labelling (Törnros et al., 2016).

3 Expected Outcome

We believe that the presented new way to attribute a value to an integrated path way may enable us to:

Assign a Personal Value to Spaces. The allocation of aggregated personal values to urban green spaces, points of interest, and other route factors provide valuable information for city planners and architects to reflect on which spaces are valuable to the citizens. By analysing the GPS data systematically it may be possible to detect universal patterns and properties of places and urban

structures of commonly high (or low) value and to analyse this under consideration of additional covariates. Making the personal, cultural, and ethical values measurable and quantifiable may improve the ability for decision makers to include these “soft” factors in decision making.

Allocate Conflicting Spaces. One finding of this analysis could also be to find “conflicting spaces” in which the value of a space (the willingness to make a detour) varies greatly. These interesting cases may be studied in more detail. If available an analysis of the background of the users (gender, age, level of education, mean of transportation, etc.) in correlation to the detour may provide insight into societal and cultural differences. This may help understand conflicts, which may arise when planning new construction projects.

Study Different Weather Conditions. As the choice of detour depends strongly on meteorological conditions (e. g., shady routes in hot weather, covered routes during rain falls, etc.), the meteorological conditions should be tracked and analysed at the same time. This allows a distinction of the value of a path depending on meteorological conditions.

Study Different Cultures. In the long run, this study can be repeated in different regions of the world to enable a top-down detection of structures (e. g., forest, meadows, cities, monuments, lakes, etc.), which are evaluated differently in different cultures. The detected differences in cultural preferences are a possible starting point for a more detailed bottom-up analysis of these patterns, which may include user surveys and interviews. Note that a global data set would be needed for such an analysis.

Weighting Routing Applications. We expect that the derived information can be used to derive personal weights required by applications for healthy or green routing.

4 Challenges

The described method of using detours as a measure for personal value has the goal to describe platials rather than spaces. It seeks to include social, cultural, and personal values by making them measurable. However, are there certain parameters, uncertainties, and behavioural patterns that require special caution and possibly special treatment in the analysis?

Not Knowing the Way. The reason for a detour may be pleasure in the detour, but it could also be that the citizen got lost or did not know that a faster way exists. One may assume that the percentages of people who get lost are relatively small compared to the people who know the way. In order to eliminate the effect of different route choices due to very small differences in the route time that cannot be resolved by the citizen, a detour time threshold must still be set under which two ways are defined as indifferent. The value of this threshold will be varied and further discussed.

Biases by the Selection of Data. The participants in the proposed approach need to be selected carefully such that they are representative of the population. Even if the sampling design was set up with a representative sample in mind a self-selection bias might be present. The conclusion must therefore take this bias into account. Additionally, some people hardly ever take detours, e. g., as they are always in a hurry. Their personal value will not be included in this analysis leading to an additional unavoidable selection bias.

Motivation of Detours. There are various motives for taking a detour. It might be for pleasure in green or scenic landscapes, it might be for a specific functional reason (to visit a friend or a bakery), it might be to exercise or it might be for sociable reasons. All these reasons provide a certain personal value to the way. A distinction of motives requires additional information on the users as, e. g., from a user survey. We plan to use the data from the Psychogeography Project to better understand the motivation of detours and to make sure that the detour is always associated to a higher benefit. Note that a differentiation of “exercise routes” (jogging or taking a stroll during lunch break) and “destination routes” (moving from A to B) should be possible because exercise routes are typically circular routes and destination routes are typically linear routes.

5 Summary and Outlook

This paper presents an idea of how to integrate a personal, social, and cultural value to georeferenced space by systematically measuring the willingness of taking detours. Analysing which detour a person is willing to take when going from A to B provides information on the personal benefit of the integrated detour path, which must compensate the additional time needed for the detour. In contrast to previous approaches, the presented method does not depend on an active labelling of a user. The great advantage is that the social value of different places can be computed automatically and comprehensively without lacking any perceptions in a top-down approach. In a next step, it may therefore serve as a tool to detect the perception of people on different places depending on factors such as different social groups or weather conditions.

This method may be helpful for city planners, architects, social scientists, conflict researchers, and policy makers because the personal evaluation of a place of the citizens need to be known to these people. The method still encounters challenges especially regarding biases in data selection, lack of knowledge of the fastest path, and the lacking information on the motivation of people. We suggest comparing the results to actual anthropological and social-science-based approaches like interviews, surveys, observations, etc. in order to make sure that this method successfully points to the personally valuable places and to distinguish the motivation for a detour. This process of matching the top-down approach with bottom-up observation can be considered as calibration of the presented automated analysis of the value of detours. The data of the Psychogeography Project on the georeferenced mood of participants can provide information on mood changes during detours and may therefore be used as a qualitative check of the personal values of detours. A scientific collaboration with respect to environmental and health sensitive routing is planned. Asking the users of an environmental and health sensitive routing app to provide feedback on a route will provide a bottom-up observation, which can be matched with the data of this study.

Notes

1. https://www.geog.uni-heidelberg.de/gis/psychogeographie_en.html



Acknowledgements

We would like to thank the Psychogeography Project, in particular, H Tost, M Reichert, U Braun, U Ebner-Primier, A Zipf, and A Meyer-Lindenberg, for providing background information and valuable notices.

Funding

S Lautenbach was funded by the Klaus Tschira foundation as part of the Heidelberg Institute for Geoinformation Technology (HeiGIT).

ORCID

Sanam N Vardag  <https://orcid.org/0000-0003-4959-9336>
 Sven Lautenbach  <https://orcid.org/0000-0003-1825-9996>

References

- Cambridge University Press: *Cambridge dictionary*. <https://dictionary.cambridge.org/dictionary/english/detour>, 2018. Retrieved 4 August 2018
- Coscieme, Luca: *Cultural ecosystem services: the inspirational value of ecosystems in popular music*. *Ecosystem Services*, 16, 2015, 121–124. doi: 10.1016/j.ecoser.2015.10.024
- Gao, Song; Janowicz, Krzysztof; McKenzie, Grant; and Li, Linna: *Towards platial joins and buffers in place-based GIS*. *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Computational Models of Place (COMP'2013)*, 2013, 42–49. doi: 10.1145/2534848.2534856

Gliozzo, Gianfranco; Pettorelli, Nathalie; and Haklay, Mordechai (Muki): *Using crowdsourced imagery to detect cultural ecosystem services: a case study in South Wales, UK*. *Ecology and Society*, 21(3), 2016. doi: 10.5751/ES-08436-210306

Lee, Heera; Seo, Bumsuk; Koellner, Thomas; and Lautenbach, Sven: *Mapping cultural ecosystem services 2.0 – potential and shortcomings from unlabeled crowd sourced images*. *Ecological Indicators*, 2018. In press

Mellon, Jonathan and Prosser, Christopher: *Twitter and Facebook are not representative of the general population: political attitudes and demographics of British social media users*. *Research & Politics*, 4(3), 2017, 205316801772000. doi: 10.1177/2053168017720008


Oteros-Rozas, Elisa; Martin-Lopez, Berta; Fagerholm, Nora; Bieling, Claudia; and Plieninger, Tobias: *Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites*. *Ecological Indicators*, 2016. doi: 10.1016/j.ecolind.2017.02.009

Richards, Daniel R and Tunçer, Bige: *Using image recognition to automate assessment of cultural ecosystem services from social media photographs*. *Ecosystem Services*, 31, 2018, 318–325. doi: 10.1016/j.ecoser.2017.09.004

Törnros, Tobias; Dorn, Helen; Reichert, Markus; et al.: *A comparison of temporal and location-based sampling strategies for global positioning system-triggered electronic diaries*. *Geospatial Health*, 11(3), 2016, 335–341. doi: 10.4081/gh.2016.473

Yoshimura, Nobuhiko and Hiura, Tsutomu: *Demand and supply of cultural ecosystem services: use of geotagged photos to map the aesthetic value of landscapes in Hokkaido*. *Ecosystem Services*, 24, 2017, 68–78. doi: 10.1016/j.ecoser.2017.02.009

A Contribution to the Visualization of the Diversity of Places

Mathias Gröbe  and Dirk Burghardt

Institute for Cartography, TU Dresden, Germany

The availability of user-generated geographic datasets offers the possibility of exploring the phenomenon “place”. In recent years researchers started some attempts to derive more information about places from the content of geo-tagged social media platforms, e. g., Twitter and Flickr. For the communicating of the analysis results, visualization is essential, but there is less knowledge available about platial visualizations. After a short review of the literature about the extraction, approaches of analysis, and visualization of places, we will present our two methods. Based on a specific use case we will utilize density based clustering and aggregation in grid cells to demonstrate quantitative information visualization of the diversity of a place on multiple scales with pie charts.

Keywords: place; visualization; cartography; multi-scale mapping

1 Places in Research

In every-day life people use places as a representation of geographic phenomena, to communicate locations in conversations, and to orientate themselves (Scheider and Janowicz, 2014; Tuan, 1977; Winter and Freksa, 2012). A place can refer to either material or immaterial entities (McKenzie and Adams, 2017) and in general describe a location together with a set of attached meanings (Cresswell, 2009). Goodchild and Li (2016) as well as Goodchild (2015) describe the possibilities of a place based analysis platform, which allows to use the potential of user-generated data. This circumstance is reinforced by the fact that there is a need for new adapted methods for the new data in a variety of fields (e. g., social life, health, transport, science) as also formulated by Tsou (2015). Another possibility is the recording of places by means of empirical studies, which can provide better results but are much more complex, as the example of Poplin (2017) shows.

Researchers have already estimated the extent of places and compared it with administrative boundaries based on multiple data sources (Hollenstein and Purves, 2010; Jaffe et al., 2006; Li and Goodchild, 2012). Sub-places and the aspects of diversity inside a place has not been in focus of research yet. It would be enlightening to uncover the things behind a place – the diversity – and allow a deeper insight. With the help of a use case and sample visualizations we would like to address this issue.

1.1 Reviewing Approaches to Derive and Visualize Places

Until now there has been a small number of practice-oriented publications that have dealt with places and used different visual analysis methods. The article by Hollenstein and Purves (2010) deals with data from Flickr and how to describe city cores, which are often implicit places. For the visualization of the raw data the authors use a point map and the kernel density estimation. In the end they provide a volume surface with contour lines for the place names in London as quantitative visualization.

M Gröbe and D Burghardt (2018): *A Contribution to the Visualization of the Diversity of Places*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 67–73

<https://doi.org/10.5281/zenodo.1472751>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

The visualization of places from spatial footprints by Li and Goodchild (2012) is similar and situated in Paris (France). Larger areas are delimited, but not more finely graded and a certain hierarchy in the places is considered. A weighted kernel density estimation was used by Chen and Shaw (2016) to derive the spatial extent. The figures provide only the estimated extent in combination with the boundaries of different US states. The juxtaposition of thematic regions of out different sources by McKenzie and Adams (2017) leads to examples for regions with Mexican, American, or Chinese Restaurants in Los Angeles, based on social media posts from Yik Yak, Twitter, and Instagram. Moreover, temporal dynamics of spatial content were explored as well as differences between the social media platforms.

It is remarkable that all publications try to delimit places. For sure it is of interest to compare administrative borders and mental boundaries, but in the end one should not only concentrate on delimiting but rather seeing the different places in context to each other. Furthermore, it should be noted that the very thin data basis, especially with regard to the usage habits of geo-tagged social media compared to the number of people visiting these places every day, does not allow a generally valid, meaningful demarcation of places. For the visualization they often used a kernel density estimation, but provide no quantitative information in the figures and only a few or no background information.

1.2 Research Questions

In the reviewed publications, the visualization was not the main content, rather a means to an end. Nevertheless, for the communication of place-related behaviours and the presentation of platial data, appropriate representations are essential. That leads us to the question: How to visualize platial information and their diversity in a quantitative way? In addition, we would like to discuss whether it is worthwhile to make a visual distinction between places. It might be more relevant to look at the diversity as categorical data within a place or region instead of just concentrating on the demarcation of one place. In addition, what occurred on different scales? Nowadays, map users are used to zooming and getting more information. Using the following case study we want to present the possibilities and assess their practical effects on the visual result. Ultimately, this could be significantly influenced by the conclusions drawn by the user.

2 Visualizing Platial Information

2.1 Methods for Visualization

Point Map. The nominal point symbol map shows data at their origin location. The symbols are variable in shape, orientation, and colour (Kraak and Ormeling, 2009). It allows the map user to identify the specific position of each platial information and informs about the spatial distribution. This basic method was used by Hollenstein and Purves (2010).

Statistical Surface Derived by Kernel Density Estimation. The kernel density estimation interpolates from some sample data (in most cases points) to an area (O'Sullivan and Unwin, 2010). The result is a density surface that can be visualized and indicates that these values are probably present in the displayed area. The result depends on the chosen kernel shape, a possible value for weighting, the radius or distance in which points are taken into account for the calculation, and the raster resolution. The result is a grid with an estimated continuous distribution of values, which then can be represented by means of colour sequences. Hollenstein and Purves (2010), Li and Goodchild (2012), Chen and Shaw (2016), and McKenzie and Adams (2017), e. g., used the kernel density estimation.

Clustering and Diagrams. Spatial clustering is the process of grouping objects into classes. The DBSCAN algorithm (Ester et al., 1996) generates clusters based on a defined distance and rejects points which are too far away from the cluster. It is a density-based method, which regards clusters as regions of a high number of objects. For the presentation of the quantitative information pie charts can be used, which show the numerical proportion where as well as the size of the diagram scales with the number of items. The result is called point diagram map (Kraak and Ormeling, 2009) but in this case, it refers more to invisible areas.

Grid-Based Aggregation and Micro Diagrams. Micro diagrams (Gröbe and Burghardt, 2017) are a visualization method, which uses small diagrams based on aggregation to show the distribution of

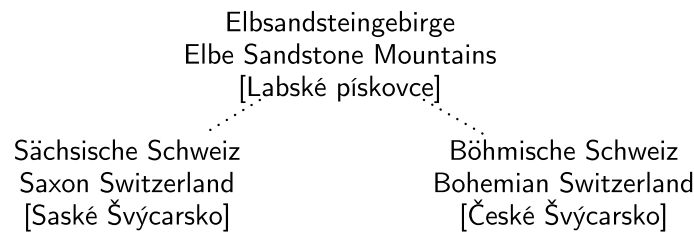


Figure 1: Structure of the sample region: The Elbsandsteingebirge consists of the Sächsische Schweiz (situated in Germany) and the Bömische Schweiz (situated in the Czech Republic). The Czech names are given in brackets.

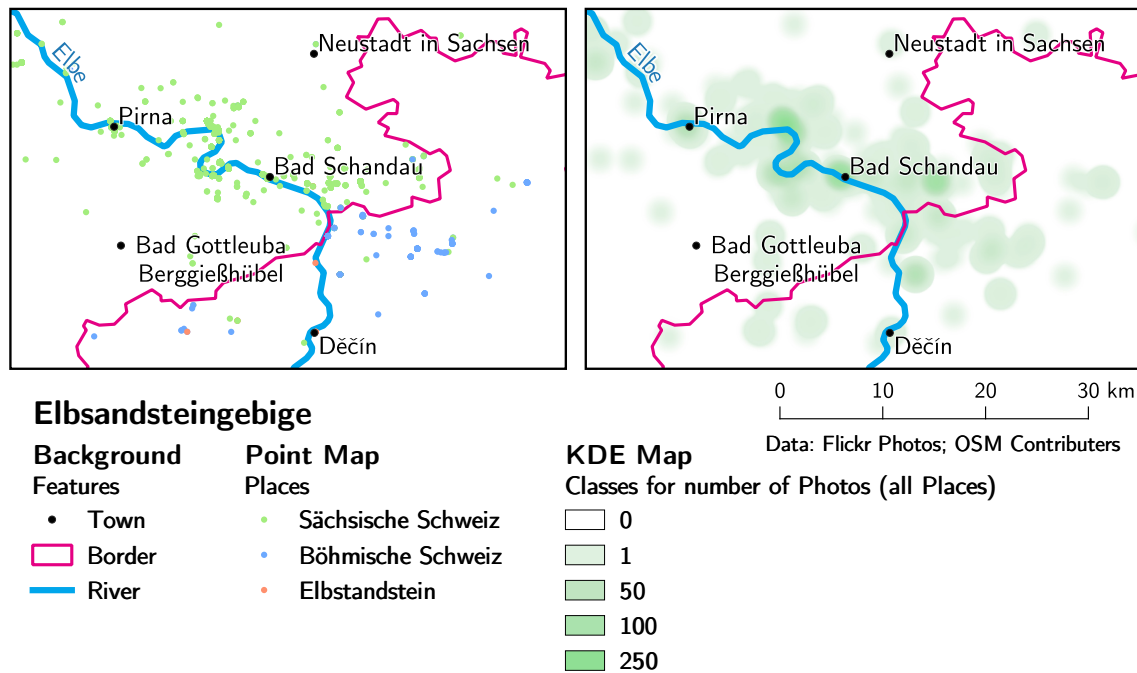


Figure 2: Common visualizations of place data. The left map shows the points with the different categories. The right map shows a kernel density estimation over all points. Data basis © OpenStreetMap contributors (cf. www.openstreetmap.org/copyright)

categorized dense point data. The aggregation can be performed in a grid that is suitable for the map scale. Nevertheless, the grid width is decisive for the result of the aggregation and should be chosen carefully. The pie charts visualize the number of aggregated points by their size using standardized diameters. This leads to a more detailed overview and can be seen as a kind of proportional symbol grid maps (Kraak and Ormeling, 2009).

2.2 Use Case

Sample Place Region. For the sample visualizations of a place we decided on the Elbsandsteingebirge. This is a landscape in Germany and the Czech Republic, which is further differentiated into Sächsische Schweiz (German part) and Böhmisches Schweiz (Czech part) by the situation in the respective countries. In our experience, data from photo platforms have proven to be suitable for analysis, since the authors usually refer to their real situation in the descriptions of their photos and do not comment remotely on what is happening at another location. To extract the data, we searched for the landscape names in German and Czech in the image descriptions as well as the tags and merged the counterparts so that the regions are combined despite the language border. Figures 2 and 3 show the visualization of this information, while Figure 1 clarifies the hierarchy of terms in the different languages.

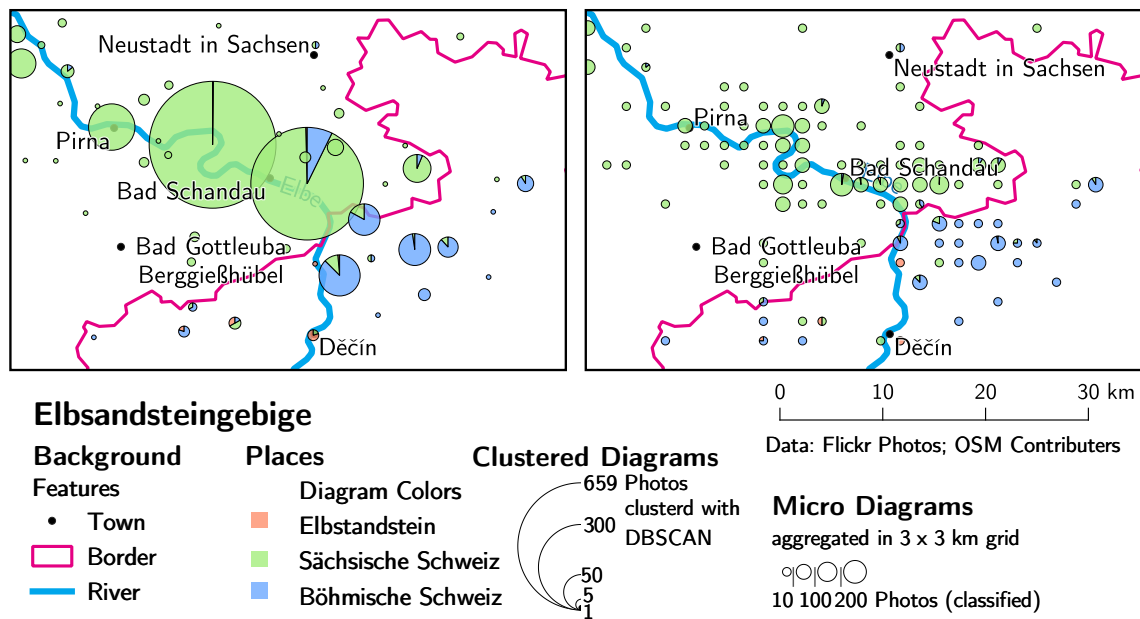


Figure 3: The two new approaches to describe the place Elbsandsteingebirge with a map. The left map uses area-proportional diagrams. The right map makes use of micro diagrams. Data basis © OpenStreetMap contributors (cf. www.openstreetmap.org/copyright)

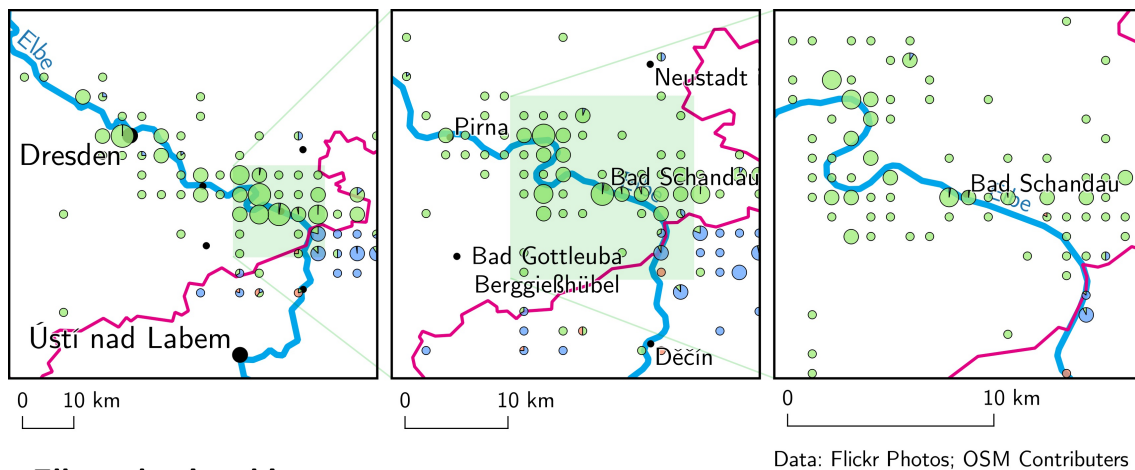
Samples for the Visualization of Platial Information. The border between Germany and the Czech Republic is shown as a background information in Figures 2 and 3. It is a 500 year old cultural border that divides the Elbsandsteingebirge into the two parts. An important landmark is the Elbe river, which has created a canyon through the mountain range. The examined landscape can roughly be defined between the cities Pirna in the west, Děčín in the east, Neustadt in the north and Bad Gottleuba-Berggießhübel in the south.

The maps in Figure 2 show the basic qualitative visualization of platial information that have been used in the reviewed papers. For the three terms different colours are used in the left map, while it was not possible in the map on the right side using the Kernel Density Method to distinguish categories inside a place. In Figure 3 maps have been created using the two new methods for visualizing platial information. For the map on the left side the DBSCAN clustering was used with a distance of 2.5 kilometres and a minimal cluster size of one point. Again, the colour shows the used terms. For the creation of the micro diagrams a grid with a cell size of 3 kilometres was used. The provided information about the number of items behind the size of the diagrams was classified, which allows for an optimal usage of the available space for the visualization.

Visualizing Platial Information on Different Scales. The visualizations shown so far all had the same scale. In the following section, the effect of the scale on visualization should be demonstrated. Figure 4 shows an overview over the sample region on the left side and detail on the right side. The scale of the map in the middle is identical with the scale of the Figures 2 and 3. The three different scales change the map extent as well – it becomes clear that the landscape is also present in the surrounding area. While the left map provides the overview, the other two maps provide an increasing number of details. As a result, the map user would probably draw different conclusions about the region Elbsandsteingebirge from each map. For the visualization the micro diagram method was used with a grid size adapted to the scale.

3 Discussion

The point map in Figure 2 offers a good first qualitative overview and is simple to create. Due to the high number of points and the missing qualitative information, clustering effects are unfavourable. In addition, points overlay each other and hide possible other points of another colour, which can distort



Elbsandsteingebirge

Background

Features

- Town
- City
- Border
- River

Micro Diagrams

Places

- Sächsische Schweiz
- Böhmsche Schweiz
- Elbstandstein

Micro Diagrams

aggregated in grid

- ○ ○ ○ ○
- 10 100 200 Photos (classified)

Figure 4: Visualization of information about the Elbsandsteingebirge at three different scales using area micro diagrams to show the diversity in the place. The extent of the right map is shown in light green in the left map. The visualization demonstrates how changes of the scale and the map extent can influence the visual analysis of a place. Data basis © OpenStreetMap contributors (cf. www.openstreetmap.org/copyright)

the distribution. This effect can be reduced by a random plot order of the points.

The kernel density estimation is a useful method for the production of statistical surfaces. In regard to the comparison of a place extent with its administrative extent it is for sure a good approach. It can describe well the unsharp nature of place, but it can only visualize more in a qualitative than a quantitative way showing the number of photos overall. Furthermore, it should be noted that the interpolated surfaces are difficult to compare and can only visualize one category. The diagrams on the left map in Figure 3 provide more overview information and may be simpler to read, but they also show the diverse categories. The micro diagrams on the right side make it possible to visualize more details while the clustering directs the view to the main content.

If we compare the maps in Figure 4, we can clearly see that by changing the scale it is possible to highlight the regions that are more important (more often tagged), but which cannot be distinguished more precisely. The left map clearly allows to distinguish between Sächsische Schweiz and Böhmsche Schweiz. Clustering can aggregate quickly across this boundary like in Figure 3 on the left side, while it is clearly visible in the micro diagrams grid on the right side. They are very useful for identifying the diverse important locations in a place.

From some points of view the shown figures are still expandable: the adjustment of the parameters for the aggregation or the kernel density estimation could lead to other results and are not optimized. Another grid width or aggregation distance can already produce another result. Quite apart from the fact that there are of course further possibilities to evaluate the used data set and further to analyse it.

The presented visualization using micro diagrams addresses mostly experts and not average users, because the maps are very detailed and have to be interpreted. Figure 4 demonstrates with the maps in different scales that a fixed delimitation of the place could be difficult. In each map you could define your own boundaries and hide the variety of landscapes of the Elbsandsteingebirge with its different subregions. At least these maps show the problem of defining a meaningful border for a place. Sure, there are some mountains of the Sächsische Schweiz visible from Dresden, but is the place really there? Choosing the extent and the scale for the analysis can be a very important decision, which can have a significant impact on the result. In addition it would be nice to have a interactive visualization to explore the place. This is more intuitive and offers a wide range of additional possibilities, e. g.,

changing the clustering distance or the visualization method of the micro diagrams. As a result, a map reader may conclude other results and perceives the place in another way.


4 Conclusions

The presented approaches for visualization of the diversity of places and the contained information have all benefits for different purposes. The micro diagram method is probably most suitable for a detailed visualization of sub-places and other facts inside places, e. g., different groups of people. A kernel density estimation is good to derive a possible area for one place and clustering to provide visualizations concentrated on a few important aspects. When visualizing a place, scale and extent should be taken into account. Furthermore, the analysis should be performed on several scales to take into account the effects of spatial aggregation effects and other scale dependent effects.

Acknowledgements

Thanks to our colleague Dr.-Ing. Alexander Dunkel for providing the Flickr data for the sample.

ORCID

Mathias Gröbe  <https://orcid.org/0000-0001-9849-8676>

References

- Chen, Jiaoli and Shaw, Shih-Lung: *Representing the spatial extent of places based on Flickr photos with a representativeness-weighted kernel density estimation*. Proceedings of the 9th International Conference on Geographic Information Science (GIScience), 2016. doi: 10.1007/978-3-319-45738-3_9
- Cresswell, Tim: *Place*. In: Thrift, Nigel and Kitchin, Rob (eds.), *International encyclopedia of human geography*, Oxford: Elsevier, 2009, vol. 8. 169–177
- Ester, Martin; Kriegel, Hans-Peter; Sander, Jörg; and Xu, Xiaowei: *A density-based algorithm for discovering clusters in large spatial databases with noise*. Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD), 1996
- Goodchild, Michael F: *Space, place and health*. *Annals of GIS*, 21(2), 2015, 97–100. doi: 10.1080/19475683.2015.1007895
- Goodchild, Michael F and Li, Linna: *Formalizing space and place*. Proceedings of the 1st colloque international du CIST, 2016, 177–183
- Gröbe, Mathias and Burghardt, Dirk: *Micro diagrams: a multi-scale approach for mapping large categorised point datasets*. Proceedings of the 20th AGILE Conference on Geo-Information Science (AGILE), 2017
- Hollenstein, Livia and Purves, Ross: *Exploring place through user-generated content: using Flickr tags to describe city cores*. *Journal of Spatial Information Science*, 1, 2010, 21–48. doi: 10.5311/JOSIS.2010.1.3
- Jaffe, Alexandar; Naaman, Mor; Tassa, Tamir; and Davis, Marc: *Generating summaries and visualization for large collections of geo-referenced photographs*. Proceedings of the 8th ACM international workshop on Multimedia information retrieval (MIR), 2006, 89–98. doi: 10.1145/1178677.1178692
- Kraak, Menno-Jan and Ormeling, Ferjan: *Cartography. Visualization of geospatial data*. London, UK: Routledge, 2009, 3rd edn.
- Li, Linna and Goodchild, Michael F: *Constructing places from spatial footprints*. Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information, 2012, 15–21. doi: 10.1145/2442952.2442956

McKenzie, Grant and Adams, Benjamin: *Juxtaposing thematic regions derived from spatial and platial user-generated content*. Proceedings of the 13th International Conference on Spatial Information Theory (COSIT 2017), 2017. doi: 10.4230/lipics.cosit.2017.20

O'Sullivan, David and Unwin, David J: *Geographic information analysis*. Hoboken, NJ, USA: John Wiley & Sons, 2010. doi: 10.1002/9780470549094

Poplin, Alenka: *Cartographies of fuzziness: mapping places and emotions*. The Cartographic Journal, 54(4), 2017, 291–300. doi: 10.1080/00087041.2017.1420020

Scheider, Simon and Janowicz, Krisztof: *Place reference systems. A constructive activity model of reference to places*. Applied Ontology, 9(2), 2014, 97–127. doi: 10.3233/AO-140134

Tsou, Ming-Hsiang: *Research challenges and opportunities in mapping social media and big data*. Cartography and Geographic Information Science, 42, 2015, 70–74. doi: 10.1080/15230406.2015.1059251

Tuan, Yi-Fu: *Space and place. The perspective of experience*. Minneapolis, MN: University of Minnesota Press, 1977

Winter, Stephan and Freksa, Christian: *Approaching the notion of place by contrast*. Journal of Spatial Information Science, 5, 2012, 31–50. doi: 10.5311/JOSIS.2012.5.90

Data Mining of Network Events With Space-Time Cube Application

Viktor Putrenko¹, Nataliia Pashynska², and Sergii Nazarenko¹

¹*World Data Center for Geoinformatics and Sustainable Development,
Igor Sikorsky Kyiv Polytechnic Institute, Ukraine*

²*Taras Shevchenko National University of Kyiv, Ukraine*

At the moment, there is a problem of studying the detailed relationship of social phenomena and their attachment to the territory. The simple means of conducting such research is getting harder due to the large amount of data that needs to be processed. It is suggested to use the methodology of space-time cube construction as one of the types of data mining with spatial-temporal distribution. The use of this method on the example of information analysis from subscribers of one of the major mobile operator networks allows to carry out statistical analysis and to detect statistically significant spatio-temporal clusters in the data that can be used during data structuring in order to provide safety and react quickly to hazardous situations.

Keywords: data mining; spatial-temporal cube; national security

1 Introduction

Starting from 2013, Ukraine's national security becomes more and more important. Every day, citizens of Ukraine meet with the most diverse threats of natural, technological, social, and military character. Dangerous processes, extreme events, catastrophes, virtual, real terrorism, etc. – these are exactly the things to which public authorities need to react practically every day.

The modern world is extremely rich in the most diverse information, the vast arrays of which people collect, store, analyze, and on the basis of it tries to make forecasts and predictions. This is especially true for information that directly and indirectly affects human security. Equally important information for government bodies, whether public or private institutions, the correct analysis of which allows you to take a step in the right (progressive) direction. But operating with big data are very problematic.

When it comes to eliminating the consequences of a certain disaster, where the bill goes, sometimes, for a minute, is needed a quick and balanced solution. For example, international statistics show that the number of rescued after the earthquake directly depends on the beginning of rescue operations. If the saviours arrive in the earthquake zone in the first three hours, they can save up to 90% of the survivors, after six hours only 50%. Only the means of rapid response can reduce the number of victims by 20–30% (ESRI, 2017a; Rak et al., 2017). That is why it is important to get information and to give an answer immediately after an emergency and not only after a while and, moreover, to prevent an emergency in advance, which can lead to a significant number of victims.

V Putrenko, N Pashynska, and S Nazarenko (2018): *Data Mining of Network Events With Space-Time Cube Application*. In: R Westerholt, F-B Mocnik, and A Zipf (eds.), *Proceedings of the 1st Workshop on Platial Analysis (PLATIAL'18)*, pp. 75–82

<https://doi.org/10.5281/zenodo.1472753>



First Workshop on Platial Analysis (PLATIAL'18)
Heidelberg, Germany; 20–21 September 2018

Copyright © by the authors. Licensed under Creative Commons Attribution 4.0 License.

2 Background

Recently, security issues become increasingly significant due to the increasing number of threats to ordinary people and the region or the country as a whole. One of the options for solving this issue is to study, analyze, and forecast the event by building a spatial-temporal cube. For the first time, the use of the space-time cube was proposed by Hägerstrand (1970) in the early 70's, whose possibilities he described in his work "What about people in regional science?". Despite the active development of geographic information systems (GIS), its use was limited. Only in the 2000s there are works on the use of the spatial-temporal cube in GIS. In the works of this period, new possibilities for using the spatial-temporal cube were presented using GIS, including earthquake surveys (Andrienko and Andrienko, 1999; Andrienko et al., 2003; Gatalsky et al., 2004; Kraak, 2003).

The next steps in using the spatial-temporal cube method were its application in the intelligent analysis of data of a variety of nature: crime analysis, infrastructure studies, animal behaviour analysis, human motion visualization, dependence studies on weather conditions changes over time, etc. (Baas, 2013; Cheng and Adepeju, 2013; Gonçalves et al., 2014; Hurcilava et al., 2013; Shapiro and Hall, 2017; Yusof et al., 2014). In the field of data mining, Ukraine is widely known for the Institute for Applied System Analysis NTUU "KPI", World Data Center for Geoinformatics and Sustainable Development (Petrenko, 2008; Putrenko, 2017; Putrenko and Pashynska, 2017; Zgurovsky et al., 2013; Zgurovsky and Pankratova, 2005).

3 Goal and Tasks

The goal of the work is to analyze the spatial-temporal regularities in the distribution of events in the Vodafone network based on the use of the methodology of space-time cube construction. Among the tasks are to study the methodology of using the space-time cube for the data mining of spatial-temporal data; the study of the application peculiarities of the space-time cube construction method for the analysis of space-time series of data generated by users of Vodafone telecommunication network; the use of building space-time cube for distribution analysis of spatial and temporal patterns of mobile data for the purpose of emergency response to natural and social emergencies.

4 Spatial-Time Cube

Spatial-time cubes are a three-dimensional visualization technology designed to simultaneously represent spatial and temporal characteristics of motion. Accordingly, trajectory points are displayed in three-dimensional space, where the vertical axis usually expresses time (Peuquet, 1994).

In the early 70's Hägerstrand (1970) proposed the use of a graphical approach to reflecting time as an addition to spatial measurements. He developed a three-dimensional diagram as a spatio-temporal cube, which allows to visually explore space-time events and interactions of processes. The cube's base reflects a flat geographic dimension, and the cube's height is time. Initially, the tool was designed to manually reproduce graphics.

The use of the space-time cube requires spatial and temporal data, for the purpose of analyzing certain events. Examples of such events include earthquakes, road accidents, cases of disease, and the observation of rare animals (Gatalsky et al., 2004). Hägerstrand proposed to apply the space-time cube to the data on the motion of objects on the changes of spatial sites with an anchor to time. In this paper, the authors propose to apply the concept of Hägerstrand to another type of data, namely, to analyze network events.

Using the spatial-temporal cube makes it possible to answer three questions of Puke (Andrienko et al., 2003), supplied to spatial-temporal data (ESRI, 2017b):

- "what, depending on when and where": description of objects or a set of objects that are present in a certain place or a set of locations for a certain time or time interval;
- "where, depending on when and what": a description of the location or set of locations occupied by a particular object or set of objects at a specific time or time interval; and

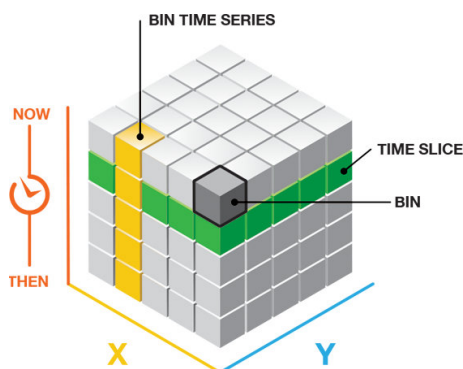


Figure 1: Bin display in the space-time cube (Source: ESRI ArcMap ESRI 2017b)

- “when, depending on where and what”: a description of a specific time or interval of time when a particular object or set of objects occupied a particular place or set of locations.

The research uses the data provided by Vodafone, which has spatial and temporal bindings as well as some attribute information. The processed database has 1.5 million network events from the most diverse devices and from different subscribers. All events are concentrated practically in the western regions of Ukraine and has geographic coordinate system WGS84. The location coordinates are recorded in decimal degrees and determine the location of the nearest mobile communication station through which the message or call was sent. All data has been collected between July and August 2017. The authors use a set of tools for in-depth analysis of spatial and temporal regularities in the software ArcMap 10.5. This toolkit contains tools for analyzing data distribution and identifying patterns in the context of space-time.

The dataset structure has a combined set of attributes that characterize the nature of the communication event, location, feature of calls and devices, as well as subscriber preferences. The description of network events is the event type, which is divided into incoming-outgoing calls, SMS, and Internet usage. The location is described by the direction and coordinates of the signal receiving station. The peculiarities of network events include the tariff plan, the category of numbers, the amount of Internet traffic, the cost of use, and the type of device. Personal preferences of the client are presented in the form of three attributes describing the interests located in the first, second, and third place for the subscriber. Examples of such preferences are the categories of science, culture, tourism, travel, football, etc.

The creation of a spatio-temporal cube takes place by arranging point data of events in space and time in the form of a cubic structure. The base unit of the cube is the bin of space-time (Figure 1), which counts the number of points in time and each location using the Mann-Kendall statistics (ESRI, 2017b).

The spatial-temporal cube consists of rows, columns, and time steps, which together form the total number of bins in the cube. The rows and columns correspond to the placement of objects in the latitude and longitude plane. The cube height corresponds to the time period. If an event occurred for a certain period of time, it will be fixed in a certain bin with spatial-temporal characteristics.

5 Mann-Kendall Test

As input objects there can be only point classes that describe the events that have taken place. Such events may include network events, emergencies, trade operations, and other events that are time-consuming and space-based. In order to obtain valid data of distances calculations, rectangular coordinate systems with corresponding projections are used.

An important part of the operation of the tool is the analytical operations over the data bins used during the simulation. The basic set of operations is the definition of the general trend of data, which is calculated on the basis of time series. Using trend analysis allows to determine the positive or negative trends in the number of events. The trend analysis is based on Mann-Kendall's statistics.

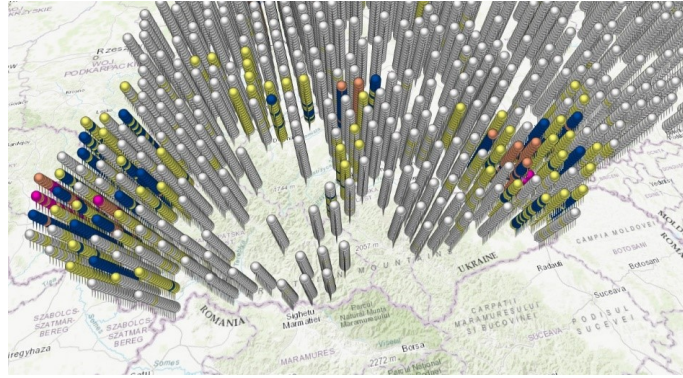


Figure 2: Three-dimensional visualization of space-time cube for western regions of Ukraine

The non-parametric Mann-Kendall test is commonly employed to detect monotonic trends in series of data. The null hypothesis H_0 is that the data come from a population with independent realizations and are identically distributed. The alternative hypothesis H_A is that the data follow a monotonic trend. The Mann-Kendall test statistic is calculated by

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(X_j - X_k) \quad \text{where} \quad \text{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

The mean of S is $E[S] = 0$ and the variance σ^2 is given by

$$\sigma^2 = \frac{1}{18} \left(n(n-1)(2n+5) - \sum_{j=1}^p t_j(t_j-1)(2t_j+5) \right)$$

where p is the number of the tied groups in the data set and t_j is the number of data points in the j -th tied group. The statistic S is approximately normal distributed provided that the following Z -transformation is employed:

$$Z = \begin{cases} \frac{S-1}{\sigma} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sigma} & \text{if } S < 0 \end{cases}$$

The statistic S is closely related to Kendall's τ as given by

$$\tau = \frac{S}{D} \quad \text{where} \quad D = \left(\frac{1}{2}n(n-1) - \frac{1}{2} \sum_{j=1}^p t_j(t_j-1) \right)^{1/2} \left(\frac{1}{2}n(n-1) \right)^{1/2}$$

The resulting Vodafone data set is in the time period from June 1, 2017 to August 31, 2017. For the convenience of analysis, the authors used a 5-day time step. As a result, the tool built a cube with a height of 19 bins (Figure 2).

6 Hot Spots Analysis

Tool analysis of hot spots determines trends in the cluster of density of points (calculations) or fields of sums in a cube. The categories of cold and hot spots include the following characteristics (Anselin, 1995): new, consistent, growing, constant, declining, sporadic, and fluctuating historical (Figure 3).

The Hot Spot method calculates a statistic for each event in the data set. The final values of p (probability) and z -estimates (standard deviations) indicate in which region of the space clustered events with high or low values exist (Andrienko et al., 2003). The method analyzes each event in the context of the neighbouring geography of events. To be a statistically significant hot spot, the event

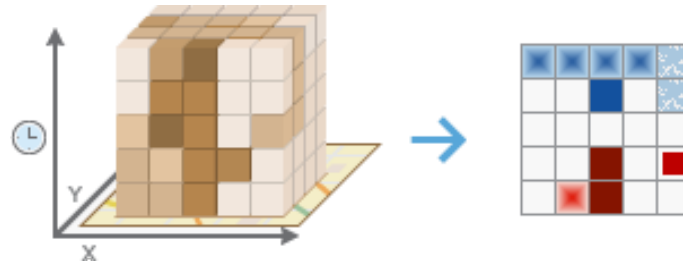


Figure 3: Transformation of a space-time cube for analysis through hot spots

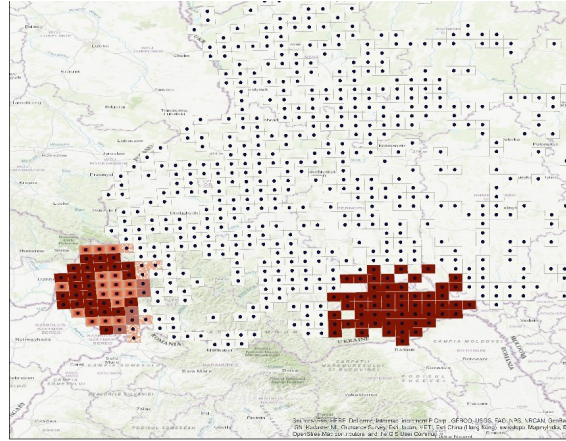


Figure 4: Map occurrence of hot spots and the resulting space-time analysis cube

must have a high value and needs to be surrounded by other approaches with also high values. Hot dots statistics uses the following formulae:

$$G_j^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \cdot \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}} \quad \text{where} \quad \bar{X} = \frac{1}{n} \sum_{j=1}^n x_j \quad \text{and} \quad S = \sqrt{\frac{1}{n} \sum_{j=1}^n x_j^2 - \bar{x}^2}.$$

Here x_j is the attributive value for the event j ; $w_{i,j}$ is the spatial weight between the events i and j ; n is the total number of events; \bar{X} is the mean of the arithmetic values of the course; and S is the dispersion.

The statistical value G_i gives each object in the set its own z-score. If the z-score has a positive value, then the probability of the intensity of the clustering of hot spots increases, which is proportional to the size of the positive estimate. Negative z-values are directly proportional to the intensity of clustering of low values and correspond to cold points.

Output objects are added to the table of contents and represent a summary of the spatial-temporal analysis for all the analyzed locations. In addition to creating a class of output objects, the summary analysis results are recorded in the results window (Figure 4).

7 Local Outlier Analysis

The analysis tools group includes the local outlier analysis tool, which allows to identify significant statistical data in both space and time. To determine statistically significant data outliers, the Anselin's Local Moran I statistics is the used statistic option, which calculates the value of each bin relative to its neighbours (Figure 5).

Based on the calculation of z and p-values of Anselin's Local Moran I statistics, each time series receives the coded value of belonging to a particular cluster with the corresponding statistical value.

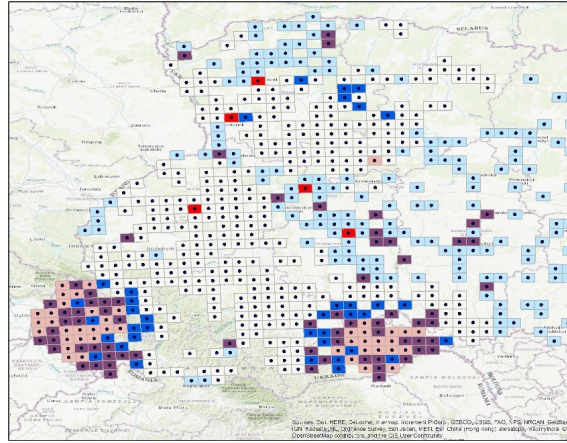


Figure 5: Map of local outliers, created as a result of the analysis of the space-time cube

The Local Moran's I statistic of spatial association is given as

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \cdot \sum_{j=1, j \neq i}^n \omega_{i,j} (x_j - \bar{X})$$

where x_j is an attribute for feature i ; \bar{X} is the mean of the corresponding attribute; $\omega_{i,j}$ is the spatial weight between feature i and j ; n the total number of features; and

$$S_i^2 = \frac{1}{n-1} \sum_{j=1, j \neq i}^n (x_j - \bar{X})^2.$$

The Z_{I_i} -score for the statistics are computed as

$$Z_{I_i} = \frac{I_i - E[I_i]}{\sqrt{V[I_i]}} \quad \text{where} \quad E[I_i] = \frac{1}{n-1} \sum_{j=1, j \neq i}^n \omega_{i,j} \quad \text{and} \quad V[I_i] = E[I_i^2] - E[I_i]^2$$

The presence of positive evaluations for I is a certificate that is adjacent to objects with similar values that may be part of a cluster. Negative values indicate the difference between the estimates of the object and its neighbours. In all cases, the value of p for the object must be so small such that the cluster is determined to be statistically significant.


To determine the belonging of the bin to the clusters, the rules of the conceptualization of spatial relationships are first defined, which determine the belonging of the bin to one of the clusters. Further, the values of bins are estimated in proximity to the centre of the cluster. Bins with high values of local emissions contain abnormal changes in the behaviour of users, which may have a different nature, both positive and negative. Together with the use of classifiers and social news dissemination channels, they can be identified and transmitted to relevant government agencies and services.


8 Conclusions


The toolkit for building a spatial-temporal cube provides a convenient visual interface for data mining of big data. The use of the spatial-temporal cube is practically possible in virtually all areas where it is necessary to analyze the behaviour of objects and events occurring with the change of location in space and time.

An example of the use of spatial-temporal analysis of data for events in mobile networks, e. g., of the Vodafone network, makes it possible to use the data more effectively, primarily for security purposes, which will be useful to governmental organizations for the rapid detection or prevention of dangerous situations, such as terrorism, extraordinary events, catastrophes, etc. In the future, using the spatial-temporal cube based on the data of mobile operators, it is possible to analyze the statistical emissions in the activity of subscribers in calls or connecting people to the Internet with an anchorage of a certain territory, which will allow to identify certain anomalies and respond accordingly.

ORCID

Viktor Putrenko  <https://orcid.org/0000-0002-0239-9241>

Nataliia Pashynska  <https://orcid.org/0000-0002-0133-688X>

Sergii Nazarenko  <https://orcid.org/0000-0003-3367-5875>

References

Andrienko, Gennady and Andrienko, Natalia: *Interactive maps for visual data exploration*. International Journal Geographical Information Science, 13(4), 1999, 355–374. doi: 10.1080/136588199241247

Andrienko, Natalia; Andrienko, Gennady; and Gatalsky, Peter: *Exploratory spatio-temporal visualization: an analytical review*. Journal of Visual Languages and Computing, 14(6), 2003, 503–541. doi: 10.1016/S1045-926X(03)00046-6

Anselin, Luc: *Local indicators of spatial association – LISA*. Geographical Analysis, 27(2), 1995, 93–115

Baas, Maarten: *Space-time cube analysis of animal behaviour*. Master's thesis, Wageningen University, 2013

Cheng, Tao and Adepeju, Monsuru: *Detecting emerging space-time crime patterns by prospective STSS*. Proceedings of the 12th International Conference on GeoComputation, 2013, 281–285

ESRI: *ArcGIS analysis workflows for public safety*. <https://community.esri.com/docs/DOC-10345-arcgis-analysis-workflows-for-public-safety>, 2017a. Retrieved 6 August 2018

— *A quick tour of geoprocessing tool references*. <http://desktop.arcgis.com/en/arcmap/10.3/tools/>, 2017b. Retrieved 6 August 2018

Gatalsky, Peter; Andrienko, Natalia; and Andrienko, Gennady: *Interaction analysis of event data using space-time cube*. Proceedings of the Eighth International Conference on Information Visualization, 2004, 145–152. doi: 10.1109/IV.2004.1320137

Gonçalves, Tiago; Afonso, Ana Paula; and Martins, Bruno: *Visualizing human trajectories: comparing space-time cubes and static maps*. Proceedings of the 28th International BCS Human Computer Interaction Conference (HCI), 2014

Hurcilava, Khurtsilava Kostiantyn; Beleckij, Boris Alexandrovich; and Bespalov, Vyacheslav Petrovich: *On the question of space-time gis and some of their applications*. In: *Proceedings of the Conference on Decision Support Systems, Theory and Practice (DSS)*. 2013, 52–55

Hägerstrand, Torsten: *What about people in regional science?* Papers in Regional Science, 24, 1970, 7–24. doi: 10.1111/j.1435-5597.1970.tb01464.x

Kraak, Menno-Jan: *The space-time cube revisited from a geovisualization perspective*. In: *Proceedings of the 21st International Cartographic Conference (ICC)*. 2003, 1988–1996

Petrenko, Anatoly Ivanovich: *Grid and data mining for intellectual data processing*. System research and information technology, 4, 2008, 97–110

Peuquet, Donna J: *It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems*. Annals of the Association of American Geographers, 84(3), 1994, 441–461

Putrenko, Viktor: *Data mining of relationship in crowdsourcing projects and social activities of citizens*. Proceedings of the 1st Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2017, 1060–1065. doi: 10.1109/UKRCON.2017.8100413

Putrenko, Viktor and Pashynska, Natalia: *Application of geoinformation modeling tools for intelligent data analysis of fire hazards*. Bulletin of National Technical University “KhPI”, Series New Solutions in Modern Technologies, 7, 2017, 156–163

Rak, Jacek; Bay, John; Kotenko, Leonard, Igor Popyack; Skormin, Victor; and Szczypiorski, Krzysztof: *Computer network security*. Proceedings of the 7th International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security (MMM-ACNS), 2017

Shapiro, Ben Rydal and Hall, Rogers P: *Interaction geography in a museum*. Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI), 2017, 2076–2083. doi: 10.1145/3027063.3053146

Yusof, Norhakim; Zurita-Milla, Raul; Kraak, Menno-Jan; and Retsios, Bas: *Mining frequent spatio-temporal patterns in wind speed and direction*. Proceedings of the 17th AGILE Conference on Geo-Information Science, 2014, 143–161. doi: 10.1007/978-3-319-03611-3_9

Zgurovsky, Mikhail Zacharovich; Boldak, Andrey Aleksandrovich; and Yefremov, Kostiantyn Viktorovich: *Intelligent analysis and the systemic adjustment of scientific data in interdisciplinary research*. Cybernetics and Systems Analysis, 49(4), 2013, 541–552. doi: 10.1007/s10559-013-9539-7

Zgurovsky, Mikhail Zacharovich and Pankratova, Natalia Dmytrievna: *System analysis: problems, methodology, applications*. Kyiv, Ukraine: Naukova Dumka, 2005