JOURNAL OF SPATIAL INFORMATION SCIENCE
Number N (YYYY), pp. xx–yy
doi:10.5311/JOSIS.YYYY.II.NNN

RESEARCH ARTICLE

Bananas and tangerines spilled on streets

Martin Fleischmann

Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool, United Kingdom

Anastassia Vybornova

NEtworks, Data and Society (NERDS), Computer Science Department, IT University of Copenhagen, Denmark

Received: December 24, 2015; returned: February 25, 2016; revised: July 13, 2016; accepted: September 5, 2016.

Abstract: 1-2 sentences basic introduction into field. 2-3 sentences more detailed background. 1 sentence clearly stating the general problem being addressed by this particular study. 1 sentence summarizing the main result ("here we show"). 2-3 sentences explaining what the main result reveals/adds. 1-2 sentences to put results into more general context. Optional - if accessibility is enhanced by this: 2-3 sentences to provide broader perspective.

Keywords: street networks, blocks, urban form, shape analysis, urban morphology, urban morphometrics, routing

1 Introduction

- Importance of street networks for urban analysis - talk about availability of data, different use cases from transport to morphology to ... - try to illustrate the wide applicability so we can then base the claims about the importance of the issue on top of it - general motivation - framework of urban data science. why everyone would benefit from having this issue solved. cite arcaute on 'recent advances, lobo on 'urban science'. also: alessandretti 2020, louail 2015, barthelemy books (morphogenesis 2018; spatial networks 2022) - Data need to look different for transport than for morphology ans why it matters - Networks vs polygons enclosed by networks (blocks? negative space? we need to pin down the terminology we want to use) - Problem description - Each network comes with a different detail and generated "blocks" are not always what they seem to be but sometimes are

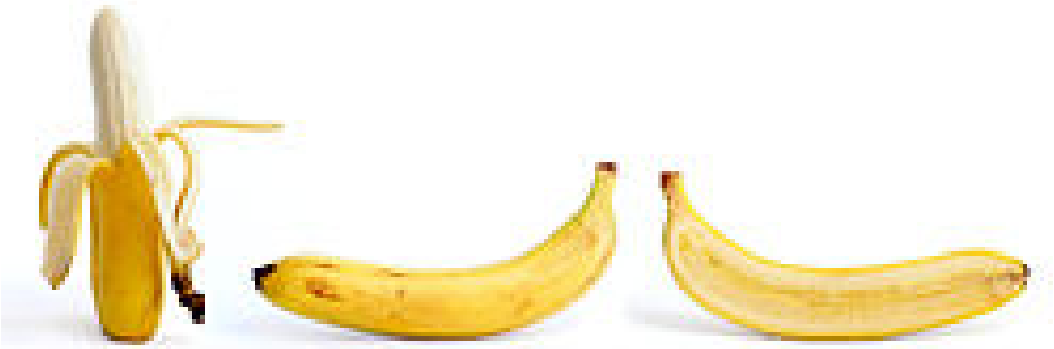

Figure 1: Banana

an artifact of transport-focused geometry - cite cardillo, geisberger, morer (computational costs), maybe venerandi 2016?; vanegas paper on actually *simulating* these spaces - Examples - other authors complaining about the issue, without having solved it yet (e.g. best paper ever [1]); grippa 2018; peponis 2007 merges these into urban blocks (replacing by center lines) - include morphometric literature here - mention ‘momepy.Blocks‘ algorithm that attemps to go around the issue in a specific way (but does not solve it) - (fleischmann, porta, dibble, etc.) diet 2018 on planar map classification. sharifi on urban forms. - description/terminology: cf. hermosilla 2014 ’UBRSA’; see strano 2012 for power law of "land cells" (spaces surrounded by street segments); most recent: shpuza 2011, 2017, 2022 (how to get the PDF...). circular compactness - inspired by louf; see also more recent barthelemy 2017 with the same figures; - summary of what happens in this paper - ’towards an automated detection of bananas’; method inspired by louf and barthelemy; tried out on 150 cities across the world

2 Method

This is method. And the figure 1 shows the banana.

- select sample of urban areas (FUA) - fetch the data from OSM - polygonize the network - measure shape charactersitics - TODO: measure initally more than Reock (get a sample from ESDA) - there is a conceptual backbone to this - we know that the artifacts are either small (small intersections) polygons or can be large but then they are very narrow (in between dual carriegaway) - we need a shape metric that captures this relationship - identify optimal measurements - plots that help us visually detect a cluster of artifacts - derivation of 1-dimensional index - from Roeck and area we can derive one value from which distribution we can identify a cut-off value for artifact/non-artifact polygons - cut-off value detection - exploration of geographical variation - differences between cities and continents - open tools, open data, open code with full reproducibility

3 Results

- area vs shape plots - use all cases together and show multiple shape indices - Reock as an optimal index (?) [I think it will be the optimal one but we need to verify that] - 1-dimensional index formula (if we use Reock it is the one from the banana notebook) - shape-index plots with cut-off values - plots based on geographical location - distributions, Reock-area scatters - describe the differences - formalise the detection workflow

4 Discussion

How could this be used?

how to move forward? (sneak preview of google summer of code) - the simplification problem can be seen as a problem of the elimination of banana

incorporate further data (ideas: directionality; street names; angles; land use; ...) use network formalism: on dual approach (intersections = edges): jiang 2004, yang 2022, rosvall/sneppen; barthelemy paper on shortest path shape

end with a call to action & 'towards open urban data science'

Acknowledgments

To be added. Remember to include ESRC/ATI funding covering initial experiments.

References

[1] VYBORNOVA, A., CUNHA, T., GÜHNEMANN, A., AND SZELL, M. Automated detection of missing links in bicycle networks. *Geographical Analysis* (2022).