Individuals' contribution to cells: a cellular automata approach for simulation of collaboratively mapped areas

Jamal Jokar Arsanjani
University of Heidelberg/Institute of Geography
Berlinerstrasse 48, Heidelberg, Germany
Jokar.arsanjani@geog.uni-heidelberg.de

Lukas Loos
University of Heidelberg/Institute of Geography
Berlinerstrasse 48, Heidelberg, Germany
Lukas.loss@geog.uni-heidelberg.de

Marco Helbich
University of Heidelberg/Institute of Geography
Berlinerstrasse 48, Heidelberg, Germany
Marco.helbich@geog.uni-heidelberg.de

Mohammed Bakillah
University of Heidelberg/Institute of Geography
Berlinerstrasse 48, Heidelberg, Germany
Mohammed.bakillah@geog.uni-heidelberg.de

Abstract

Geolocated user-generated content (GUGC) appeared as a result of Web 2.0 technologies. A huge number of applications have been launched to gather geo-located information from volunteered people. The GUGC is being constantly and increasingly generated, which demands studying the spatio-temporal pattern of contributors' activities. The main objective of this paper is to monitor the mapping behaviour of the volunteers within the lifetime of a well-known project, OpenStreetMap. The spatio-temporal evolution of contributions in this project is simulated by embedding the location of mapped areas into a lattice of cells. Then, a cellular automata model is designed and

1. Introduction

The evolution of Web 2.0 led to the creation of a number of online mapping websites for diverse purposes providing geolocated user-generated contents (GUGC) [2]. GUGC provides new and updated information in a variety of fields such as routing optimization, environmental monitoring, and recreational designation [1]. One of the major advantageous of GUGC is to provide some data that are totally new and up to date [3,7]. Neis & Zipf [4] discovered that OpenStreetMap's (OSM) total route length already exceeds that of TomTom at a rate of 27%, particularly in eastern Germany. A number of collaborative mapping projects have been launched, which aim at collecting geodata. OSM has been a pioneer project due attracting the most contributions so far exceeding 1,110,000 users by April 2013. Therefore, analysing the emergence and evolution of this project can answer some arising questions about the future of individuals' contributions.

Cellular automata (CA) modeling

CA is part of a family of rule-based simulation techniques based on customized and simple rules that can simulate the behaviour of a phenomenon across time and space [6]. It has been broadly utilized in a variety of fields for studying dynamic phenomena e.g., urban sprawl, fire spread, disease dissemination, and social sciences. Theoretically, CA consists of a set of cells projected in a lattice indicating actors. The cells emerge in one of a number of states, for instance "infected" and "not-infected", and may change their state concurrently at each discrete time stamp depending on their neighbours' states [5]. This manuscript aims to consider the mapping process in the OSM project as a dynamic phenomenon.

2. Material and methods

Study area

For this study, Heidelberg in Germany is selected. Justifications for choosing this study area are twofold: firstly, Germany is a pioneer country in having a significant quantity of OSM contributions; and secondly, the picked area of interest is composed of urban and rural areas together.

Data pre-processing

The data used in this investigation is the OSM nodes extracted from the OSM-Dump file in August 2012. A cellular space representing the nodes' footprints is projected at the spatial resolution of 50 m covering 2500 m^2 with 4988 pixels in total. Six timestamps from 22 July 2006 until 22 July 2012 with one-year sequences are determined. The cellular network contains either non-contributed cells (values of 0) or contributed cells (values of 1).

Methodology

Firstly, the OSM nodes from the study area are collected and embedded on a lattice of cells. Next, a quantitative analysis on the amount of contributions is carried out. Transition rules are defined as the crucial component of CA, which is an iterative process to realize whether the associated transition rules are suitable. The output that correlated best with the actual map is verified to be applied for the prediction process. Secondly, a CA model is configured and certain transition rules are learnt after a training process.

3. Results

Spatio-temporal analysis

The spatio-temporal analysis of contributions is illustrated in Figure 1, which represents a non-substantial rate of contribution within the first year and an increasing trend of new contributions and users from 2009 on. By 2012 most of the cells were contributed with only a small portion of cells left to be contributed.

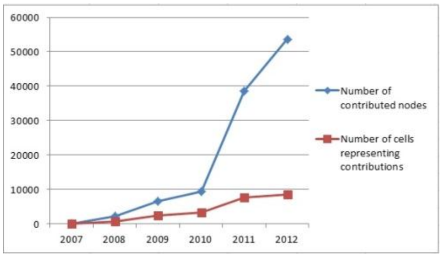

Figure 1: Increasing trend of contributions to OSM nodes

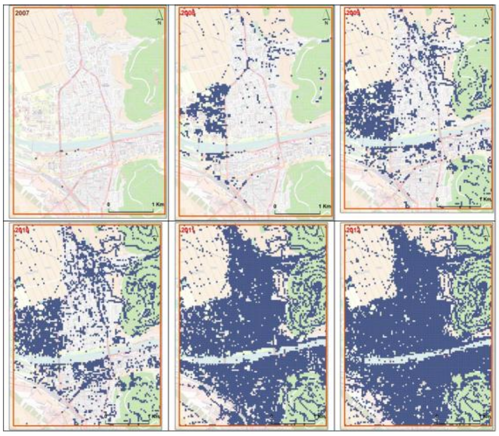

Figure 2: Spatio-temporal representation of contributions

Model implementation and prediction

In order to configure the simulation process, three timestamps (2009, 2010 and 2011) are selected to train the model and calibrate it. By using them, future contributions are predicted and compared with the actual values. A kappa index analysis is carried out to evaluate the outcomes against reality. Accordingly the most optimal transition rules and contiguity filter are recognized and applied for the prediction process. Consequently, the expected contributions for 2013 and 2014 are predicted as shown in Figure 3. Visual analysis indicates that the simulated contributions will cover the whole urban fabric. Nonetheless, the north-western part of the study area, which contains farming lands, will not receive further attention through individuals' contribution.

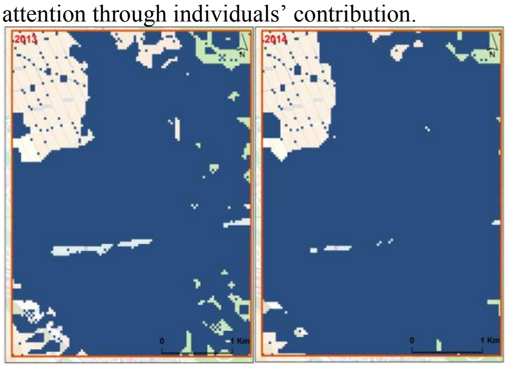

Figure 3: The predicted quantity and location of future contributions in the study area

4. Discussion and conclusions

GIS functions provide several opportunities to monitor, model and simulate dynamic phenomena across time and space. So far, a large number of spatially explicit techniques have been developed and applied on dynamic phenomena to study their behaviour and also predict their future actions. In this study, social activity into a public mapping call (mapping in OSM) is spatially and temporally monitored and simulated. In this study, by embedding every single contribution into a lattice of cells, this social activity was printed on a spatial domain. Besides, cellular automata model was implemented to properly learn the behaviour of individuals in order to fit it into a simulation paradigm. This study investigated the collaboratively mapped environment from the first mapped object from 2006 to mid-2012 at six timestamps.

Based on the trend of contributions, the pattern of contributions seems to be related to socio-economic and physical indicators. The earliest and the most prevalent contributions were received on urban fabrics and primarily on the campus area of the Heidelberg University, where education is taking place and highly educated people commute there. Instead, farming lands and water bodies receive contributions later than other features. This implies that socio-economic variables should be coupled as well.

References

[1] Castelein, W., Grus, L., Crompvoets, J., & Bregt, A. (2010). A characterization of Volunteered Geographic Information. 13th AGILE International Conference on Geographic Information Science 2010, Guimarães, Portugal.

[2] Haklay, M., Singleton, A., Parker, C., 2008. Web mapping 2.0: the Neogeography of the geoweb. Geography Compass 2 (6), 2011–2039.

[3] Mooney, P. & Corcoran, P. (2012). Characteristics of Heavily Edited Objects in OpenStreetMap. Future Internet, 4(1), 285-305.

[4] Neis, P., & Zipf, A. (2012). Analyzing the Contributor Activity of a Volunteered Geographic Information Project – The Case of OpenStreetMap, 5.

[5] Spicer, V., Reid, A. a., Ginther, J., Seifi, H., & Dabbaghian, V. (2012). Bars on blocks: A cellular automata model of crime and liquor licensed establishment density. Computers, Environment and Urban Systems, 36(5), 412–422.

[6] Wolfram, S. (1984). Cellular automata as models of complexity. Nature, 311(5985), 419–424.

[7] Zielstra, D. and Zipf, A., (2010). A comparative study of proprietary geodata and volunteered geographic information for Germany. In: Proceedings of 13th AGILE international conference on GIS, 10-14 May 2010, Guimaraes, Portugal.