Genetic algorithms for map generalization

Sergio Solano Pinedo, Marco Moreno Ibarra and Miguel Torres Ruiz

Centro de Investigación en Computación
Instituto Politécnico Nacional
México, DF.
ssolano_a13@sagitario.cic.ipn.mx
{mtorres,marcomoreno}@cic.ipn.mx

1 Motivation

A map is usually thought of as a tool (describing a part of the world) for the user, but when a map is poorly designed instead of a tool it becomes an obstacle for the user [1]. A common error when designing maps is the amount of information shown. In a map, inefficiency in information communication is proportional to the amount of irrelevant information according to a user query [2]. When a user asks for specific data in the spatial information context, it is common to find information overload [3], that is, the user gets a very difficult to read map, with much useless data; a map content must be appropriate for the map scale and purpose.

A user also has to look at a map time enough for him to discern the information within the map [4]. This observation time varies depending on the user's familiarity and experience with maps, knowledge on the map's theme and its ability to discern visual signs on the map [1]. So, it is assumed that the better the map design is, the easier it'll be to read it.

Thereby, the motivation for this work is to create personalized maps containing familiar information (well-known places) to the user in order to improve the map readability. The generalized maps (less detailed maps) will be derived from highly detailed maps, thus leading to a spatial generalization problem. As study case, we will work with Points Of Interest (POI) from Mexico City

2 Previous works in the area

In Geographic Information Systems (GIS), the term *generalization* is used to describe a process in which a less detailed map is derived from a highly detailed one [5]. It is used to correctly pick information from a highly detailed source, avoiding most of all information overload [6]. The generalization process deals with the selection of objects and their adjustment for the less detailed map. In generalization it is vital to reduce detail [7], keeping only the most evident, meaningful attributes according to the map's scale and purpose [5]. Of course, if the map size decreases then the number of objects in the map most also decrease, otherwise the map would be cluttered [8]. Thus, generalization can be understood as a competition for space [9], where depending on scale, one must choose which objects to retain.

Usually, generalization algorithms are classified according to the type of geometry for which they have been designed (i.e., points, lines and polygons) [6]. This work focuses on POI's, so generalization for lines and areas remain out of the scope of this work. We split the generalization in different tasks (or generalization operations): selection, simplification and aggregation (or clustering) [11].

- Selection: The basic idea is that less important features should be omitted. All points preserved by these algorithms remain in their original positions [10].
- Simplification: These algorithms also choose points that remain in their original position, however the simplification is governed by geometric properties such as proximity, density, Voronoi neighbors or distribution.
- Aggregation: These algorithms aim to minimize a measure of dispersion by grouping together semantically similar and spatially close points. Some examples of these algorithms are K-means algorithm and the ISODATA algorithm.

Just as generalization operations, we can also find *Generalization Models* for point data. Their main purpose is to assign importance to objects [12] so one can segregate the less relevant ones. Some models give as result a sorted list of points, while others generate unsorted lists (or groups of points).

Besides the models, when it comes to implementation we can also find different techniques. In [8] they work with ε *-approximations* where they propose the ε value (based on the Radical Law [13]) and this represents the number of elements wanted for the generalized map. In [14] they also use the Radical Law to calculate the number of elements, then they compute the importance of each point (considering multiple attributes) in the source map and with Voronoi diagrams the pick the most relevant ones.

When it comes to *Artificial Intelligence* there is not much research regarding generalization, one work is by [15]. They work with Self Organizing Maps (SOM) to generalize lines, specifically they use SOM for streets clustering considering multiple attributes such as topology and geometry. SOM have also been used in other spatial research but no research has been made in point generalization, thus we propose to approach the generalization problem using *Genetic Algorithms*.

3 Hypothesis or research objectives

A generalization is more efficient when it considers multiple attributes of data, therefore we consider *Genetic Algorithms* a good approach to evaluate multiple attributes and discern the most outstanding ones.

The main goal of this work is to design and implement a generalization methodology that considers qualitative criteria for creating a spatial personalized representation. As secondary goal we look forward to designing a generalization methodology using *Artificial Intelligence* tools.

4 Methodology

The current state of the methodology describes the general stages that will be detailed as this thesis progress. Our priority for now is to define how to generalize points using *Genetic Algorithms* (Fig. 1), and we focus on how to qualify the map, i.e. define the *Fitness Assessment* module to evaluate each generated map. The *Fitness Assessment* block shall consider all the attributes that together will determine the relevance of the POI and simultaneously identify a solution to the generalization.

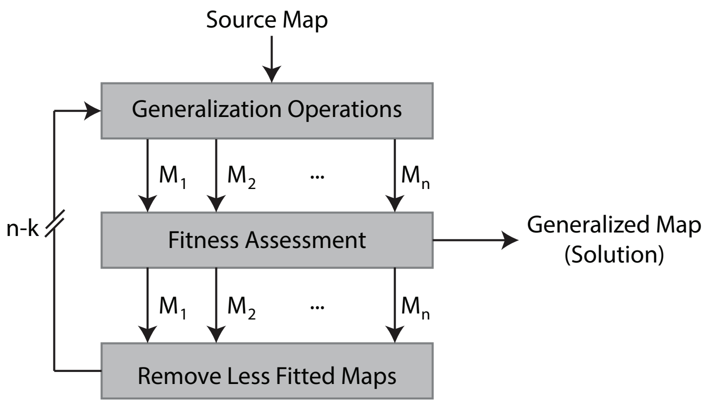

Fig. 1. Detail of the methodology.

Fig. 1 shows that the algorithm will produce a generation of n maps (individuals). Then the generation fitness will be tested, if a solution is found the algorithm will stop, otherwise the k less fitted individuals will be eliminated and replaced again in the *Generalization Operations* module.

5 State of research

We are delimiting the problem and establishing how achievable it is to face the problem of point generalization using *Genetic Algorithms*. We are also researching on how to evaluate a generalized map and have found two possibilities: an evaluation done once the generalization has ended; and an evaluation during the generalization (a sort of feedback). Conserning the first evaluation (*generalization evaluation*), and because it is a way to establish the quality of generalization for points [8], we are focusing on an evaluation based on similarity compared against the original data. In regards to the *feedback evaluation* we are planning what attributes should be simultaneously considered [14] to rate the generalization progress and we are also planning which restrictions must be considered to achieve a "good design" on the generalized map (how to state "good design"). We plan to use this evaluation within the *Genetic Algorithm*.

6 Preliminary results

Our work is still on a design phase, we have not started testing our hypothesis. So far we have no results.

Acknowledgments. The authors thank the Consejo Nacional de Ciencia y Tecnología as well as the Instituto Politécnico Nacional for the facilities provided for the realization of this research, especially the Secretaría de Investigación y Posgrado.

References

1. Sarjakoski, L. T., Nivala, A. M.: Adaptation to Context - A Way to Improve the Usability of Mobile Maps. In: L.Meng, A. Zipf, T. Reichenbacher (Eds.), Mapbased mobile services - Theories, Methods and Implementations, pp. 107-123. Springer. (2005)
2. Knöpfli, R., Communication theory and generalization. In: D. R. F. Taylor ed. Graphic communication and design in contemporary cartography, 19, New York, Wiley. (1983)
3. Doyle, J., Han, Q., Weakliam, J., Bertolotto, M., Wilson, D., Developing Non-proprietary Personalized Maps for Web and Mobile Environments. In: Y.-J. Kwon, A. Bouju, and C. Claramunt, W2GIS 2005. (2005)
4. Triglav, J., Geolocation of millennium development goals. In: GeoInformatics Magazine for Geo-IT Professionals. 2 March 2004 Vol.7. (2004)
5. Mackaness, W., Generalization of Spatial Databases, In: The Handbook of Geographic Information Science, Blackwell Publishing, (2008)
6. Zaragoz, B., Gimnez, P., Navarro, J. T., Dong, P., Ramón, A. Development of free and opensource GIS software for cartographic generalization and occupancy area calculations. Ecological Informatics, 8, 4854. (2012)
7. Goodchild, M. F. Scale in GIS: An overview. In: Geomorphology. (2011)
8. De Berg, M., Bose, P., Cheong, O., Morin, P. On simplifying dot maps. Computational Geometry, 27(1), pp. 4362. (2004)
9. Competition for map space as a paradigm for automated map design. In: Proc. GIS/LIS'90 pages 97-106. (1990)
10. Bereuter, P., Weibel, R. Generalisation of point data for mobile devices: A problem-oriented approach. the 13th Workshop on Progress in Generalisation , 18. (2010)
11. Li, Z.: Algorithmic Foundation of Multi-Scale Spatial Representation. CRC Press, Florida (2007)
12. van Kreveld, M., van Oostrum, R., Snoeyink, J. Efficient settlement selection for interactive display. In: Proc. Auto-Carto. (1997)
13. Töpfer, F., Pillewizer, W. The principles of selection. In: The Cartographic Journal 3, pp. 10-16. (1966)
14. Yan, H., Weibel, R. An algorithm for point cluster generalization based on the Voronoi diagram. Computers and Geosciences, 34(8), pp. 939954. (2008)
15. Jiang, B., Harrie, L. Selection of Streets from a Network Using Self-Organizing Maps. Transactions in GIS, 8(3), pp. 335350. (2004)