Effects of Irregular Topology in Spherical Self-Organizing Maps

Charles R. Schmidt
School of Geographical Sciences
Arizona State University
Charles.R.Schmidt@asu.edu

Sergio J. Rey
School of Geographical Sciences
Arizona State University
srey@asu.edu

André Skupin
Department of Geography
San Diego State University
skupin@mail.sdsu.edu

Abstract

We explore the effect of different topologies on properties of self-organizing maps (SOM). We suggest several diagnostics for measuring topology-induced errors in SOM and use these in a comparison of four different topologies. The results . . .

1. Introduction

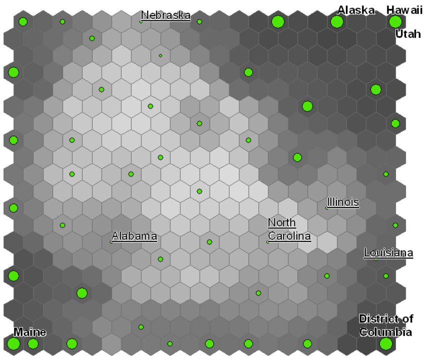
The Self-Organizing Map (SOM) is an unsupervised competitive learning process developed by Teuvo Kohonen as a technique to analyze and visualize high dimensional data sets. The applications of SOM are far reaching; Kohonen [4] provides a thorough review of the SOM literature including applications of SOM. SOM has been used in applications ranging from speech recognition and image classification to breast cancer detection and gene expression clustering. Agarwal and Skupin [10] outline the growing interest of SOM to the GISciences, and propose that the relationship between SOM and GIScience should be bidirectional. The SOM offers a powerful method for exploring and visualizing geographic data and GIScience offers a wide array of tools and methods to enable the exploration of the SOM itself. The exploration of spatial relationships has always been of great interest to geographers, and as Ritter states, the goal of SOM is "to translate *data similarities* into *spatial relationships*" [8, p. 1].

The SOM is a type of artificial neural network in which neurons are "organized" in such a way as to project the high-dimensional relationships of a set of training data onto a low-dimensional network structure. The traditional SOM uses a rectangular or hexagonal network topology [4]. These topologies create a well-known problem in SOM called the boundary or edge effect. Neurons on the boundary of the hexagonal and rectangular lattices have fewer neighbors, which reduces their ability to interact with other neurons during the self-organizing process. Using a spherical lattice has been widely suggested as a solution to the problem [8, 1, 9, 6, 12]. The use of the spherical lattice, however, does not completely overcome the boundary problem, and the choice of which spherical topology to use for the network can be difficult to make.

A regular network topology is one in which every node on the network has exactly the same number of adjacent nodes. Any topology involving an edge is irregular. Arranging our lattice on the surface of a sphere seems to be an obvious way to overcome the edge. However, there exist only five arrangements on the sphere which are completely regular; these are the five platonic solids [8, 2]. Any other arrangement of neurons on the surface of the sphere will result in an irregular topology, as not all neurons will have the same number of neighbors.

The classic method for minimizing this irregularity is to generate the spherical lattice by tessellating (subdividing) the sides of the icosahedron [6]. While this method will always result in a highly regular spherical topology, the main drawback is that the number of neurons in the network (the network size), N, grows exponentially as tessellations are applied. That results in only very coarse control over network size. Other methods for arranging neurons on the sphere allow for unlimited control over network size, but yield topologies with increased irregularity [2, 11, 6]. To date the literature has largely ignored the more irregular methods in favor of the aforementioned tessellation-based methods. A topology which yields a more flexible network size may be desirable. However, in order to address this issue of network size, we must first determine the degree to which irregularity effects the SOM.

Figure 1. Edge effects in SOM



2. Edge Effects

An intrinsic problem with SOMs in two dimensions is the so called edge effect. This is illustrated in in Figure 1, which displays a SOM trained on socioeconomic census data consisting of 32 variables for US states. The darker a neuron in the figure, the larger is the distance between the neuron and mean of the input vectors. The distance between each of the original observations (states) and the mean is represented by a graduated circle. Taken these two together it is clear that the outlier observations are pushed towards the edges of the map, while the observations that are closest to the multidimensional center of the data are assigned to central neurons on the map.

At the edge of the map, neurons have fewer neighbors which results in any observations being assigned to them having fewer competitive signals. At the same time, the edge of the neural lattice represents a true visual boundary which affects its ability to represent data similarities as spatial relationships.

One way to eliminate the edge effect is to wrap the lattice around a three-dimensional object such as a sphere or torus, thereby removing the edge entirely. The toroidal SOM was introduced by Li et al. [5], however the torus is not effective for visualization, as maps generated from a torus are not very intuitive [3, 12]. Ritter [8] describes the torus as being topologically flat and suggests that a curved topology, such as that of a sphere, may better reflect directional data. A sphere also results in a more intuitive map, since we are accustomed to looking at geographic maps based on a sphere.

Ritter [8] first introduced the spherical SOM, and several enhancements have since been suggested [1, 9, 6, 12]. A good comparison of these enhancements can be found in Wu and Takatsuka [12]. All of these methods derive their spherical structure through the tessellation of a polyhedron as originally proposed by Ritter [8]. Wu and Takatsuka [12] point out the importance of a uniform distribution on the sphere, and that it is preferable for all neurons to have an equal number of neighbors and to be equally spaced. They find generally that the tessellation method best satisfies these conditions, and specifically that the icosahedron is the best starting point [11]. Tessellation of the icosahedron results in a network of neurons, each having exactly six neighbors, save the original twelve which each have five neighbors. This is very close to the ideal structure in which every neuron would have exactly six neighbors. Wu and Takatsuka [12] prefer these this structure, because it has very low variances in both neuron spacing and neighborhood size.

Based solely on measures of neuron spacing, Wu and Takatsuka [11] dismissed the usefulness of a method proposed by Rakhmanov et al. [7] for distributing points on a sphere. Similarly Nishio et al. [6] use these variance measures to support their helix algorithm for distributing points on a sphere. However, these metrics can be misleading and comparison across topologies may not be consistent. The traditional rectangular and hexagonal topologies have no variance in neuron spacing, and the generally preferred hexagonal structure displays greater variance in neighborhood size than the rectangular structure. The torus, by comparison, would have variance in neuron spacing, yet no variance in neighborhood size. The distance between two neurons is only considered during the formation of the neural network. At this stage the spacing is significant as it plays a part in constructing the network's topology by determining neuron adjacency. However, using this measure to evaluate potential topologies for use in SOM may be misleading.

As spherical (and other alternative) topologies become increasingly more common it is necessary to investigate how the choice of topology effects the SOM. In this thesis the effect of irregularity within topologies is studied as an attempt to investigate not only the edge effect, but also to help facilitate the comparison of topologies. It is important to note that spherical topologies may not be appropriate for all applications. Removing the edge may reduce the SOM's ability to converge. As outliers are forced to interact they introduce more competition among the neurons. We would also expect outliers to occupy more space in the final map as their dissimilarity in attribute space should translate to more distant spatial relationships in the trained SOM. More research will be needed to help researches determine the most appropriate topology for their data and research objectives.

In this paper we explore the general utility of certain irregular spherical topologies beyond offering greater control over network size. We develop and test new diagnostics to measure and visualize topology-induced errors in SOM. More specifically we

1. Compare the internal heterogeneity of observations

captured by a given neuron to that neuron's first-order neighborhood size.

2. For different topologies, compare the internal heterogeneity of each neuron against a composite measure of topological regularity.

3. Develop a SOM-based visualization of the internal heterogeneity.

These diagnostics help facilitate the evaluation of both traditional and spherical SOMs. To satisfy the objective of this research, we apply these diagnostics to a series of comparable SOMs. Each SOM is trained using the same synthetic data and training parameters, but utilize different network topologies. By formally testing for difference of means and variance in the results of the diagnostics, the following questions are addressed:

1. Does the internal heterogeneity of a neuron decrease as its first-order neighborhood size, or degree, increases?

2. Is the average internal heterogeneity of a SOM higher when a more irregular topology is used?

3. Which insights, if any, can be gained from a SOM-based visualization of internal heterogeneity?

3. Methods

We train SOMs using four different topologies: *rectangular, hexagonal, geodesic sphere* and *spherical*. The spherical topology is based on a method, developed by [7], for distributing an arbitrary number of points on to the surface of a sphere. Delaunay triangulation is then applied to these points, producing a topological structure. To yield meaningful results these SOMs must be trained with comparable parameters. The literature provides many rules of thumb for training a SOM: each SOM is trained in two stages, the first of which uses a larger initial learning rate and neighborhood search radius with a small number of training steps; the second stage uses a lower initial learning rate and neighborhood search radius, but extends the length of training.
First Stage Parameters:

- Initial neighborhood search radius of 50%, which decreases during training.

- Initial learning rate of 0.04 which decreases during training.

- 100,000 training steps.

Second Stage Parameters:

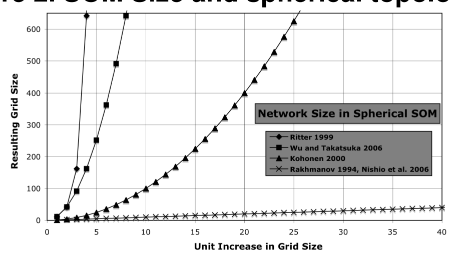
- Initial neighborhood search radius of 33%, which decreases during training.

- Initial learning rate of 0.03 which decreases during training.

- 1,000,000 training steps.

As shown in Figure 2, topologies differ in terms of achievable network size. For comparability, the network size of each SOM needs to be as close as possible. The achievable network size for the geodesic SOM is the most limiting of the topologies we test. We chose the eighth frequency geodesic sphere, which has 642 nodes, which is relatively close to the 644-node hexagonal and rectangular topologies achieved when the dimensions are set to $28x23$. Finally, the spherical topology was set to 642 nodes.

Figure 2. SOM Size and spherical topologies.



4. Results

5. Conclusions

References

[1] F. Boudjemai, P. B. Enberg, and J. G. Postaire. Surface modeling by using self organizing maps of Kohonen. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 3, pages 2418–2423 vol.3, 2003.

[2] J. M. Harris, J. L. Hirst, and M. J. Mossinghoff. *Combinatorics and graph theory*. Springer, New York, 2000.

[3] M. Ito, T. Miyoshi, and H. Masuyama. The characteristics of the torus self organizing map. In *Proceedings of 6th International Conference ON Soft Computing (IIZUKA2000)*, volume A-7-2, pages pp.239–244, Iizuka, Fukuoka, Japan, October 2000.

[4] T. Kohonen. *Self-Organizing Maps*. Springer, 3rd edition, 2000.

[5] X. Li, J. Gasteiger, and J. Zupan. On the topology distortion in self-organizing feature maps. *Biological Cybernetics*, V70(2):189–198, 1993.

[6] H. Nishio, M. Altaf-Ul-Amin, K. Kurokawa, and S. Kanaya. Spherical SOM and arrangement of neurons using helix on sphere. *IPSJ Digital Courier*, 2:133–137, March 2006.

[7] E. A. Rakhmanov, E. B. Saff, and Y. M. Zhou. Minimal discrete energy on the sphere. *Mathematical Research Letters*, 1:647–662, 1994.

[8] H. Ritter. Self-organizing maps on non-euclidean spaces. In S. Oja, E. & Kaski, editor, *Kohonen Maps*, pages 97–110. Elsevier, Amsterdam, 1999.

[9] A. Sangole and G. K. Knopf. Visualization of randomly ordered numeric data sets using spherical self-organizing feature maps. *Computers & Graphics*, 27(6):963–976, 2003.

[10] A. Skupin and P. Agarwal. Introduction: What is a self-organizing map? In P. Agarwal and A. Skupin, editors, *Self-Organizing Maps: Applications in Geographic Information Science*. Wiley, 2008.

[11] Y. Wu and M. Takatsuka. Geodesic self-organizing map. In R. F. Erbacher, J. C. Roberts, M. T. Grohn, and K. Borner, editors, *Proc. SPIE Vol. 5669*, volume 5669 of *Visualization and Data Analysis 2005*, pages 21–30. SPIE, March 2005.

[12] Y. Wu and M. Takatsuka. Spherical self-organizing map using efficient indexed geodesic data structure. *Neural Networks*, 19(6-7):900–910, 2006.