Topology Preservation in Fuzzy Self-Organizing Maps

Mohammed Khalilia[1] and Mihail Popescu[2]

[1] Computer Science
University of Missouri
Columbia, Missouri, USA
mohammed.khalilia@gmail.com
[2] Health Management and Informatics
University of Missouri
Columbia, Missouri, USA
PopescuM@missouri.edu

Abstract. One of the important properties of SOM is its topology preservation of the input data. The topographic error is one of the techniques proposed to measure how well the continuity of the map is preserved. However, this topographic error is only applicable to the crisp SOM algorithms and cannot be adapted to the fuzzy SOM (FSOM) since FSOM does not assign a unique winning neuron to the input patterns. In this paper, we propose a new technique to measure the topology preservation of the FSOM algorithms. The new measure relies on the distribution of the membership values on the map. A low topographic error is achieved when neighboring neurons share similar or same membership values to a given input pattern.

Keywords: Fuzzy self-organizing map, topology preservation, map continuity, relational data.

1 Introduction

Self-Organizing Maps (SOM) is an unsupervised neural network algorithm. SOM tries to map the *s*-dimensional input patterns to a 2-dimensional lattice, preserve the topology of the data, and cluster the neurons that represent similar input patterns, which can be visualized using a 2D or 3D map such as the Unified Distance Matrix (U-Matrix) [1]. Several formulations and modifications were proposed to the classical SOM algorithm, such as the Self-Organizing Semantic Maps [2], Ontological SOM [3], Relational Topographic Maps [4], and WEBSOM [5]. Another class of SOMs is the fuzzy SOM algorithms. The general idea of FSOM is to integrate fuzzy set theory into neural networks to give SOM the capabilities of handling uncertainly in the data. FSOM can also be divided into two categories: object FSOM [6–10] where input patterns are represented as feature vectors and the relational FSOM [11] which handles relational data.

Regardless of the type of SOM algorithm they all share one important feature that is topology preservation. Topology preservation means that neighboring data points in

M. Jamshidi et al. (eds.), *Advance Trends in Soft Computing WCSC 2013*,
Studies in Fuzziness and Soft Computing 312,
DOI: 10.1007/978-3-319-03674-8_10,

the input space are mapped to nearby neurons in the output space. Once a good mapping is established, SOM can represent the high dimensional input space in a 2-dimensional output map that preserves the topology of the input data. This in turn yields better visualization and reveals more information about the structure and the clusters presented in high dimensional input space. To ensure that SOM has established good mapping, we need to measure or quantify the goodness of SOM. Different measures are proposed to accomplish this goal, such as the quantization error and the topographic error. Those errors are widely used in SOM and while the quantization error was adapted for the object and relational FSOM [11], no formulation is yet proposed to measure the topological preservation or continuity of the map in the FSOM algorithms.

The topographic errors used in SOM are not directly applicable to FSOM due to the fact that FSOM does not assign a unique winning neuron for every object, instead every neuron is a winning a neuron of every object with a varying degree of membership. Therefore, in this work, we propose a technique to measure the topographic error in FSOM algorithms.

The reminder of the paper is organized as follows: Section 2 gives an overview of the fuzzy relational SOM. Section 3 discusses some of the well-known methods to measure the goodness of SOM. Section 4 explains a new approach to measure the topographic error in FSOM. Section 5 presents experimental results and we conclude this paper with remarks and discussion in Section 6.

2 Fuzzy Relational Self-Organizing Maps

In this section we give a very brief overview of the fuzzy relational SOM algorithm (FRSOM) [11] on which the experimental results discussed in section 5 are based on. However, the same technique for evaluating the topology preservation can be used on object FSOM or any FSOM algorithm. For a complete analysis of FRSOM the reader is referred to [11].

Given n input objects $O = \{o_1, \dots, o_n\}$ described by feature vectors $X = \{x_1, \dots, x_n\} \subset \mathbb{R}^s$ or by a relational matrix $R = [r_{jk}] = \left[\|x_j - x_k\|^2\right]$ [4,11] SOM constructs a lattice or map of c number of neurons (similar to Fig. 1a), that are connected using a neighborhood kernel, h, such the neighborhood between neuron i and j is given by

$$h_{ij} = \exp\left(\frac{-\|a_i - a_j\|^2}{2\sigma^2(t)}\right), \tag{1}$$

where a_i is the coordinate of the ith neuron in the output space (two dimensional space) and σ is a monotonically decreasing neighborhood size. Every neuron has a corresponding s-dimensional weight vector, $m = \{m_1, \dots, m_c\}$ or an n-dimensional coefficient vector in the relational algorithm. One of the goals of the classical crisp SOM algorithm is to assign every s-dimensional input signal, o_k, a winning or a best-matching unit (BMU), w_k, according to

$$w_k = \arg\min_i \|m_i - x_k\|^2 \;\; \forall 1 \le i \le c \text{ and } 1 \le k \le n \, . \quad (2)$$

Effectively, SOM assigns a full membership of o_k in neuron w_k,

$$u_{ik} = \begin{cases} 1, \text{if } w_k = i \\ 0, \text{otherwise} \end{cases} \, . \quad (3)$$

An alternative to this approach is to assign a fuzzy membership for all objects in every neuron as described in [11]. The FRSOM proposed in [11] produces fuzzy partitions $U \in M_{fcn}$ where

$$M_{fcn} = \left\{ U \in \mathbb{R}^{c\times n} \middle| \begin{array}{c} u_{ik} \in [0,1], \\ \sum_{k=1}^{n} u_{ik} > 0, \sum_{i=1}^{c} u_{ik} = 1, \\ \forall \, 1 \le i \le c \; and \; 1 \le k \le n \end{array} \right\}. \quad (4)$$

Introducing fuzzy memberships to SOM as in FRSOM adds another layer of complexity due to the fact that all neurons are winners of all objects to some degree. Thus, any error measurement made in FRSOM has to factor in all membership values of all input signals in all neurons. In [11] we showed that the quantization error in SOM can be easily adapted to the FRSOM, but this is not the case regarding the topographic error. In the next section we will briefly review two of the major SOM evaluation techniques followed by a new method to evaluate the topology preservation of FRSOM in section 4.

3 Topology Preservation in SOM

Several measures are proposed to measure the goodness of the map. Some measures, such as the quantization error, evaluate the fitness of SOM to the input data. This error calculates the average distance between the input patterns and their corresponding winning neurons [12]. Optimal map is expected to produce a smaller error, which means the input patterns are close to their winning neurons. Quantization error for SOM is shown in (5).

$$qe_c = \sum_{k=1}^{n} \|x_k - m_{w_k}\| \, . \quad (5)$$

Similarly, the FSOM quantization error is defined as [11]

$$qe_f = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^{q} \|x_k - m_{w_i}\| \, . \quad (6)$$

However, the crisp and fuzzy quantization errors in (5) and (6) may not accurately measure the topographic preservation of the map. Instead, one can quantify the relation between the codebook weight vectors and the associated neurons in the map as in the topographic product [12]. This gives a sense on how well the *s*-dimensional space is mapped to a 2-dimensional lattice [13]. A different approach is to use the topographic error.

The topographic error measures the continuity of the map or how well an input signal preserves the local continuity of the map [12]. When the first and second best-matching units to object o_k are adjacent in the map space, then o_k is said to preserve local map continuity and if they are not adjacent then there is a topological error. To evaluate the overall topology of the map the proportion of input signals for which the first and second best-matching units are not adjacent is measured (7) [12]. A lower error yields a better map and topology.

$$te_c = \sum_{k=1}^{n} adj(o_k), \tag{7}$$

where

$$adj(o_k) = \begin{cases} 1, if\ the\ first\ and\ second\ BMUs\ are\ not\ adjacent \\ 0, otherwise \end{cases}.$$

Another matric for measuring topology preservation in crisp SOM is discussed in [14]. The metric is said to be topology preserving if for any x_i, if x_j is the kth nearest neighbor of x_i, then w_j is the kth nearest neighbor of w_i.

The concept of first and second BMUs is not applicable to FSOM since every unit i is a BMU of every object o_k with a degree u_{ik}. A possible workaround is to harden the fuzzy partition produced by FSOM to find the BMU then compute the topographic error as in (7). Another approach is to consider the two neurons in which o_k has the highest membership as the first and second BMUs. However, neither of these two approaches exploits the membership grade of FSOM. Therefore, a new formulation to measure the local continuity of the map in FSOM is needed to evaluate its goodness and the topology preservation, which is the topic of the next section.

4 Topology Preservation in FRSOM

In FRSOM every neuron is a BMU of every object with a varying degree of membership. Regardless, both the crisp and fuzzy SOM should preserve the topology. Therefore, every pattern presented to FRSOM is also expected to preserve the local continuity of the map. One can consider the first and second neuron with the highest membership to o_k as best and second winning neurons, w_k and w_j. However, this flawed strategy uses only two neurons and discards all other neurons despite the fact other neurons might have high membership to o_k. Relying on two neurons can only give us a false sense of the map continuity. Consider a scenario where the first and second neurons with the highest memberships to o_k, w_k and w_j are immediate neighbors, but the neuron with the third highest membership to o_k is distant from w_k and w_j. A better approach is to use the membership values and utilize all neurons when measuring the topology preservation of FRSOM. More specifically, by looking at the differences of the membership values between the neurons and their immediate neighbors we can make a conclusion on how well the local topology of the map is preserved.

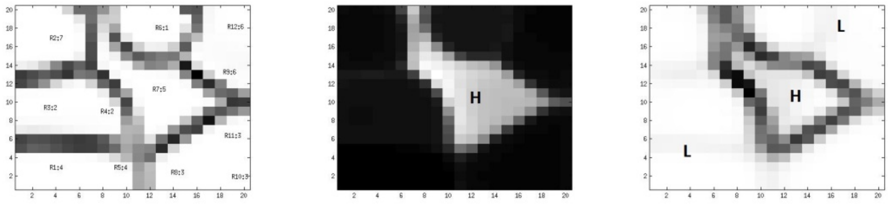
For any given object o_k in FRSOM, we expect neurons with high firing strength to o_k to be concentrated in one region (H region). Also, not all neurons have the same firing strength, as we go further away from the H region, the membership values start to diminish gradually. If the correct data topology is discovered by FRSOM, the H region corresponds to the catchment basin or part of it where o_k belongs the most. In such case, we say that o_k preserves the local continuity of the map. On the other hand, if the neurons of high membership to object o_k are scattered throughout the map or if no H region is identified then the object fails to preserve the topology of the map. For demonstration, Fig. 1a shows the topographic map for Hepta dataset [15] and Fig. 1b shows the H region for some input pattern.

In order to assess how well an object o_k preserves the local continuity of the map we first need to compute the HL-matrix. HL-matrix has the same dimensions as the topographic map and c neurons. A topology preserving HL-matrix includes two main regions, the H region which contains the neurons with high membership to object o_k and the L region containing the rest of the neurons which have low membership values to o_k, as shown in Fig. 1c. Observe that the HL-matrix of o_k represents a snapshot of the U-matrix (Fig. 1a). Adjacent neurons in regions H and L should have similar membership values to o_k. Hence, the difference in the membership values between a neuron i and its immediate neighbors $N(i)$ should be very small with exception to the bordering neurons that separate the H and L regions as shown in Fig. 1c. For a given object o_k we first compute its HL-matrix where the value at every neuron's coordinate is computed as follows

$$HL(i) = \sum_{j \in N(i)} \left| u_{ik} - u_{jk} \right| . \tag{8}$$

$HL(i)$ corresponds to the sum of differences between the membership u_{ik} and the memberships of o_k in $N(i)$. Then that difference is projected on top of the grid position of every neuron. This process is performed for every input pattern. For a small topographic error the value for every neuron $HL(i)$ should be as small as possible, which means that the neuron i and its neighbors $N(i)$ have very similar memberships to the given input pattern.

For an object to preserve the local topology it is imperative that we identify a single region labeled H. Failure in identifying a single region H will cause the topographic error to increase and possibly reaching its maximum value. This technique is stricter than the topographic error in (7). Here we want to ensure that two adjacent neurons have similar membership to o_k, which is somewhat similar to (7), but in addition we would like to ensure that o_k preserves the local continuity within a specific region of the map.



(a) (b) (c)

Fig. 1. (a) FRSOM topographic map of the Hepta dataset, (b) H region for some input pattern o_k and (c) HL-matrix of same input pattern

For more accurate evaluation of the topology preservation it is recommended that we normalize the HL-matrix as follows

$$NHL(i) = \frac{HL(i)}{\sum_{j=1}^{c} HL(j)} . \quad (9)$$

Two important reasons for this normalization: first, it sets an upper bound on the topographic error, similar to (7) the maximum error is 1. Second, normalization is crucial when comparing the topographic errors across different maps. Once the normalized HL-matrix is computed, the final topographic error of a single object o_k will depend on the neurons identified in the region labeled H. The error is simply the sum of values enclosed in the H region of the NHL-matrix (10). As the values in the H region get smaller, so does the topographic error. Meaning that adjacent neurons in the H region share similar memberships to o_k.

$$te_f(k) = \sum_{i \in H} NHL(i) . \quad (10)$$

The final topographic error of the map is computed as the average topographic error overall the objects as

$$te_f = \frac{1}{n} \sum_{k=1}^{n} te_f(k) . \quad (11)$$

Few remarks to point out about the proposed measure (11): first, the only way for a map to result in a zero topographic error is when the values in the H region are zeros. In other word, when neuron $i \in H$ and its neighbors $N(i)$ have an identical membership to o_k. Second, an HL–matrix may not contain a unique H region. In this situation the topographic error can reach its maximum, which is the sum of all values in the NHL-matrix ($te_f = 1$).

5 Experimental Results

5.1 Fuzzy Topographic Error on O3G

The overlapping three Gaussian (O3G) dataset contains three clusters of size 500 each (Fig. 2a). Clusters in O3G have larger variance which causes overlapping. We setup FRSOM with initial (σ_0), final neighborhood radius (σ_f), initial fuzzifier (q_0), final fuzzifier (q_f), map dimensions and number of epochs to be 2, 0.5, 1, 2, 15x15 and 10, respectively. The resulting topographic map is shown in Fig. 2b.

From Fig. 2c it is clear that the HL-matrix for some given pattern contains the two H and L regions, which is an indication that it preserves the local continuity of the map.

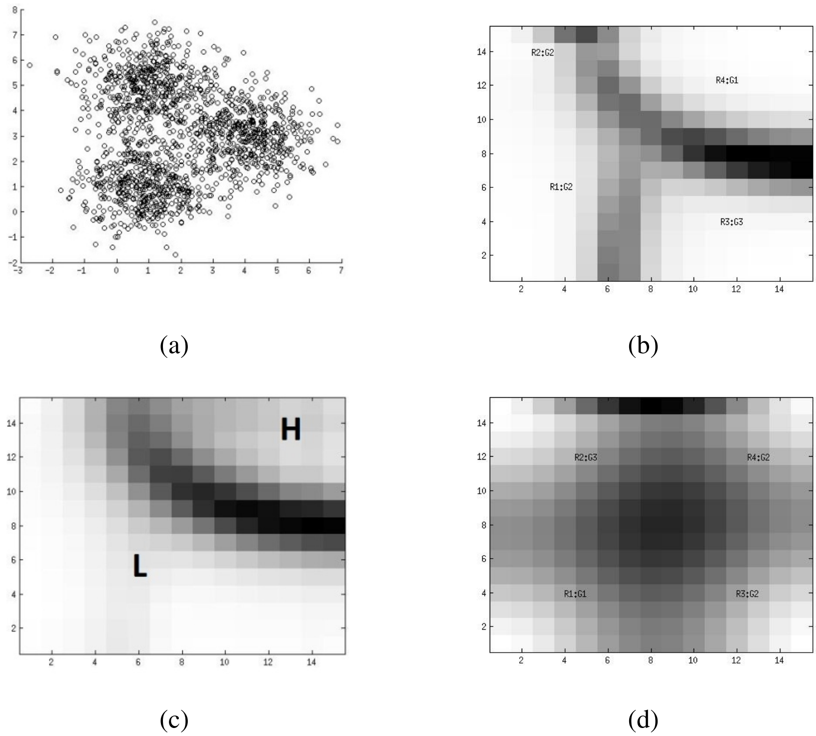
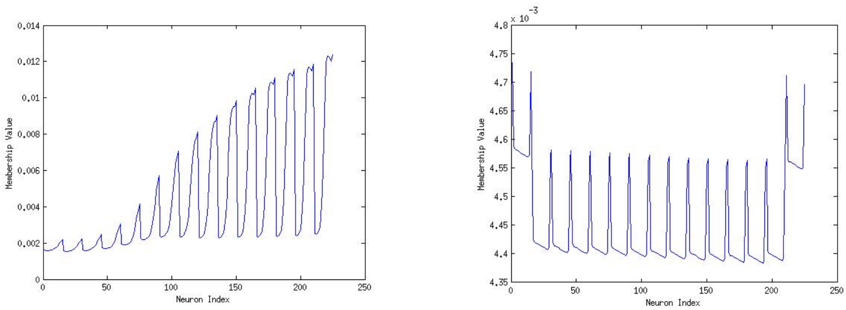


Fig. 2. (a) O3G dataset, (b) topographic map produced by FRSOM when $\sigma_0 = 2$, (c) HL-matrix for some object o_k and (d) topographic map produced by FRSOM when $\sigma_0 = 4$

In a topology preserving map, such as the one in Fig. 1c, the membership u_{ik} is expected gradually increase while approaching the H region and neurons with the highest membership should be located within the H region as demonstrated in Fig. 3a. On the other hand, a non-topology preserving map as in Fig. 2d we see a more chaotic



(a) (b)

Fig. 3. (a) Behavior of membership values of o_k in a topology preserving map (membership vs. neuron index), (b) behavior of the membership in a non-topology preserving map ($\sigma_0 = 4$)

membership values among the neurons (Fig. 3b) causing te_f to increase. It could also mean that the four regions or corners in Fig. 2d are wrapped around to form one region representing all input patterns, failing to preserve the topology.

Now, let us compare te_c and te_f for the maps in Fig. 1b and Fig. 1d. If we compute te_c for the map in Fig. 2b (Table 1), where the two neurons with the highest membership value to an input pattern are used as the first and second BMU, we find it higher than the te_c in Fig. 2d (Table 1). On the contrary, te_f has increased from 0.32 in Fig. 1b to $te_f = 1$ in Fig. 1d. In this scenario te_f reveals more information about the goodness of the map resulted from FRSOM since we probably expect Fig. 2b to be more topology preserving than Fig. 2d.

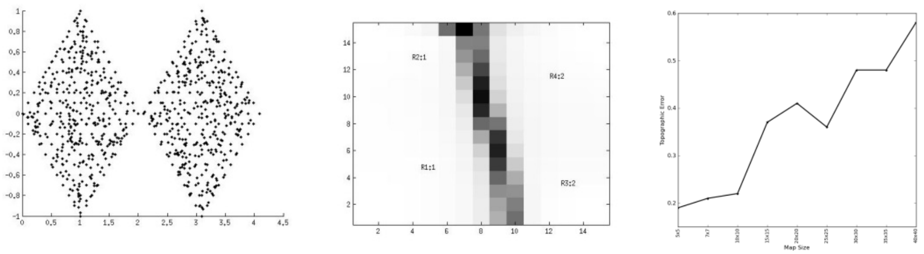
Table 1. Behaviour of te_c and te_f when varying σ_0

Map	σ_0	te_c	te_f
Fig. 2b	2	0.021 (0.006)	0.32 (0.03)
Fig. 2d	4	0.004 (0.004)	1 (0)

5.2 Fuzzy Topographic Error and Map Dimensions

In this experiment we will use the Two Diamonds dataset from the Fundamental Clustering Problem Suite (FCPS), which contains 800 data points [15] as shown in Fig. 4a. On this dataset we will show how the map dimensions can have an influence on the topographic error. Same parameters used on the O3G dataset will be used for the Two Diamonds with exception to the map dimensions which is set it be 20x20. The resulting topographic map is shown in Fig. 4b.

A smaller map of size 10x10 was also produced for the Two Diamonds dataset. It is not shown since it is very similar to the map in Fig. 4b. We found the overall topological error of the 20x20 map measured to be 0.33. As the map size increases it is likely that the H region increases which in some cases causes an increase in the membership variance among adjacent neurons. On the contrary, 10x10 map might have lower variance in the memberships among neighboring neurons in the H region and hence a lower topographic error (overall topographic error is 0.28). Overall, as the map size increased the topographic error increased (Fig. 3c). Therefore, it is important to choose a map size suitable for the dataset.



(a) (b) (c)

Fig. 4. (a) Two diamonds dataset, (b) topographic map produced by FRSOM and (c) map size vs. $\boldsymbol{te_f}$ for 5x5, 7x7, 10x10, 15z15, 20x20, 25x25, 30x30, 35x35, 40x40 map dimensions as shown along the x-axis

6 Conclusion

In this paper we presented preliminary results for measuring the topology preservation in fuzzy self-organizing maps. The newly proposed topographic error relies on the membership distribution on the map and in some sense is an extension to the crisp topographic error. The assumption is that adjacent neurons should have similar memberships to a given object o_k. In addition, we presented the HL-matrix. A topology preservation HL-matrix for a given o_k contains two regions, the H region that encompasses the neurons with high membership to o_k and the L region which contains the low membership neurons to o_k. In the results different scenarios were presented to demonstrate how the topographic error behaves when varying the map dimensions. We observed that the topographic error in FSOM tends to be higher than the standard topographic error used in SOM.

One drawback of the proposed measure is its dependence on the map dimensions. For instance, as the map dimensions or size increases so does the topographic error. To overcome this problem, one is expected to specify a map dimension that is suitable to the input dataset. The dependency of the topographic error on the SOM parameters is not necessarily a bad thing. On the contrary, a high topographic error is an indication that the map is not optimal and the parameters require tuning. However, additional experiments are needed to study the influence of other parameters such as the neighborhood size and the fuzzifier, in addition to the map dimensions, on the proposed topographic error.

Furthermore, a more theoretical approach for determining the H region is needed; contrary to the current approach of thresholding the membership values and employing image segmentation to determine the H and L regions.

References

1. Ultsch, A.: Maps for the visualization of high-dimensional data spaces. In: Proc. Workshop on Self organizing Maps (2003), http://www.informatik.uni-marburg.de/~databionics/papers/ultsch03maps.pdf
2. Ritter, H., Kohonen, T.: Self-organizing semantic maps. Biological Cybernetics 61, 241–254 (1989)
3. Havens, T.C., Keller, J.M., Popescu, M., Bezdek, J.C.: Ontological self-organizing maps for cluster visualization and functional summarization of gene products using Gene Ontology similarity measures. In: 2008 IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence), pp. 104–109. IEEE (2008), doi:10.1109/FUZZY.2008.4630351
4. Hasenfuss, A., Hammer, B.: Relational topographic maps. In: Berthold, M.R., Shawe-Taylor, J., Lavrač, N. (eds.) IDA 2007. LNCS, vol. 4723, pp. 93–105. Springer, Heidelberg (2007), http://www.springerlink.com/index/D0664R20V2L83MX5.pdf
5. Kaski, S., Honkela, T., Lagus, K., Kohonen, T.: WEBSOM – Self-organizing maps of document collections. Neurocomputing 21, 101–117 (1998)
6. Corsini, P., Lazzerini, B., Marcelloni, F.: A new fuzzy relational clustering algorithm based on the fuzzy C-means algorithm. Soft Computing 9, 439–447 (2004)
7. Pascual-Marqui, R.D., Pascual-Montano, A.D., Kochi, K., Carazo, J.M.: Smoothly distributed fuzzy c-means: a new self-organizing map. Pattern Recognition 34, 2395–2402 (2001)
8. Chi, S.-C., Kuo, R.-J., Teng, P.-W.: A fuzzy self-organizing map neural network for market segmentation of credit card. In: SMC 2000 Conference Proceedings, 2000 IEEE International Conference on Systems, Man and Cybernetics. Cybernetics Evolving to Systems, Humans, Organizations, and their Complex Interactions, vol. 5, pp. 3617–3622. IEEE (2000)
9. Mohebi, E., Sap, M.N.M.: Hybrid Kohonen Self Organizing Map for the Uncertainty Involved in Overlapping Clusters Using Simulated Annealing. In: 2009 11th International Conference on Computer Modelling and Simulation, pp. 53–58. IEEE (2009), doi:10.1109/UKSIM.2009.28
10. Bezdek, J.C., Tsao, E.C.-K., Pal, N.R.: Fuzzy Kohonen clustering networks. In: 1992 Proceedings of the IEEE International Conference on Fuzzy Systems, pp. 1035–1043. IEEE (1992), doi:10.1109/FUZZY.1992.258797
11. Khalilia, M., Popescu, M.: Fuzzy relational self-organizing maps. In: 2012 IEEE International Conference on Fuzzy Systems, pp. 1–6. IEEE (2012), doi:10.1109/FUZZ-IEEE.2012.6250833
12. Kiviluoto, K.: Topology preservation in self-organizing maps. In: Proceedings of International Conference on Neural Networks (ICNN 1996), vol. 1, pp. 294–299. IEEE (1996)
13. Arsuaga Uriarte, E., Díaz Martín, F.: Topology preservation in SOM. International Journal of Mathematical and Computer Sciences (2005), http://www.ica.luz.ve/~enava/redesn/ebooks/TopologyPreservationInSOM.pdf
14. Bezdek, J.C., Pal, N.R.: An index of topological preservation and its application to self-organizing feature maps. In: Proceedings of 1993 International Conference on Neural Networks (IJCNN 1993), Nagoya, Japan, vol. 3, pp. 2435–2440. IEEE (1993)
15. Ultsch, A.: Clustering with SOM: U* C. In: Proc. Workshop on Self- Organizing Maps, pp. 75–82 (2005), http://www.citeulike.org/group/2572/article/1304144