Persistence Homology of Collaboration Networks

Adapted from the paper "Persistent Homology of Collaboration Networks" from C. J. Carstens and K. J. Horadam

Jeff Ohana

Introduction

- Connections in social networks have different strengths (e.g., family, friends, colleagues). Strong ties are found in clusters or groups, and the weak ties connect these groups.
- We may use weighted networks to model a social network with different strengths. Weighted networks are often transformed to unweighted networks by setting a threshold weight W* and keeping only connections that are stronger than this weight.
- In this paper we discuss a recent technique from the field of computational topology called persistent homology. This method records structural properties (of a network) and their changes for a whole range of thresholds.
- However, there are some "little" inconveniences with this technique: do not inform on average path length, small world property, scale-free degree distribution and clustering coefficients.

Abstract

- We apply persistent homology to four collaboration networks
- We use persistent homology to distinguish collaboration networks from similar random networks.

Persistent homology of weighted networks

- Review on Persistence Homology

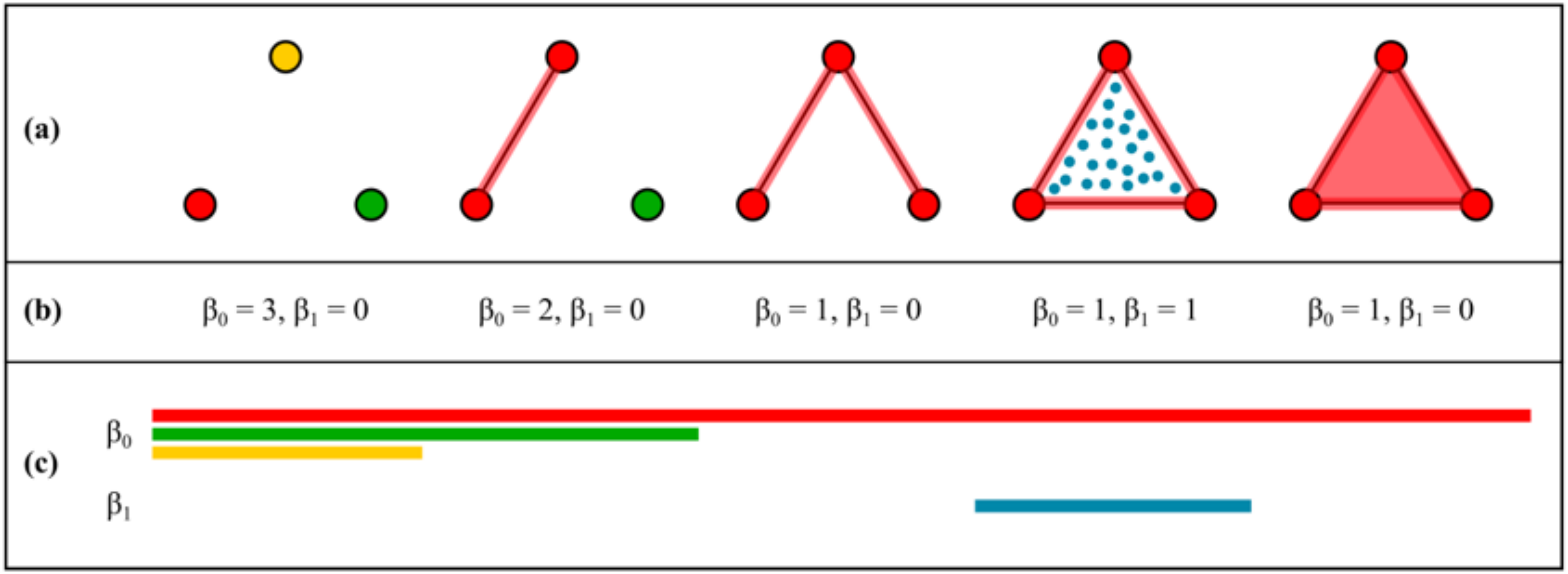


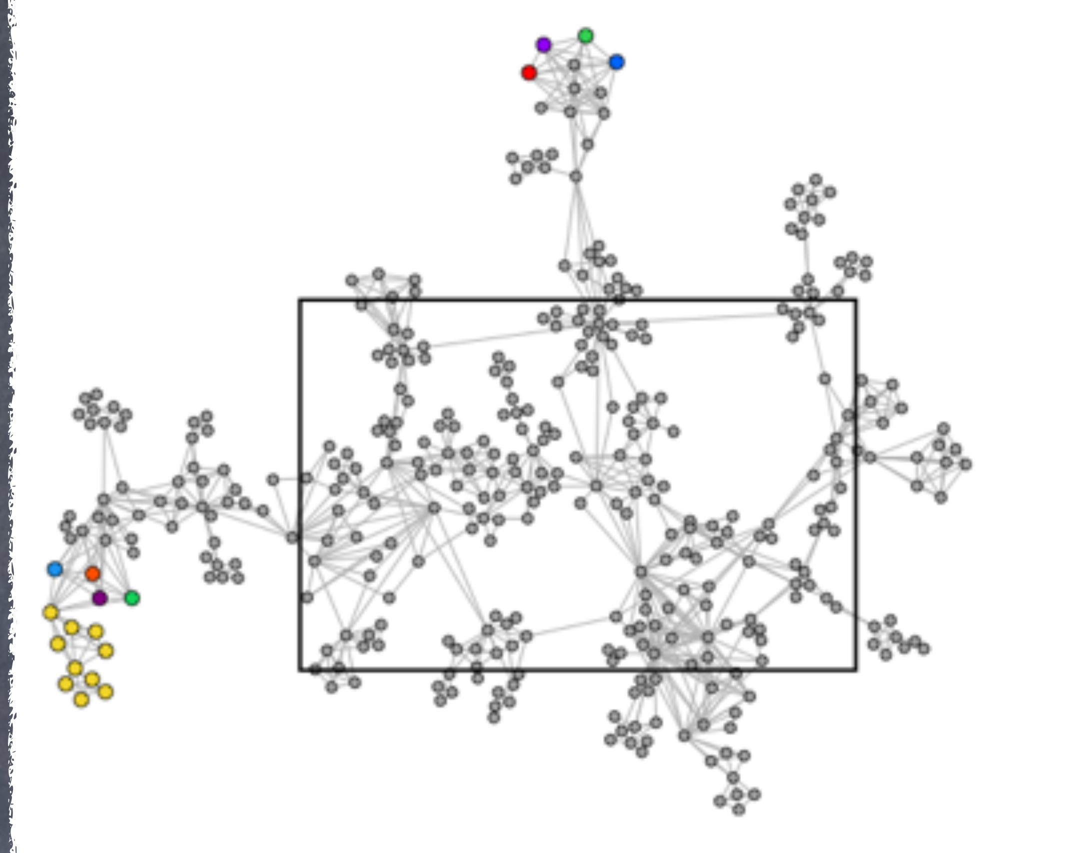
Figure 1: A filtration of a triangle **(a)**. We start with three connected components. The yellow and the green component die in step two and three, but the red component persists the whole filtration. In the fourth step a loop is born, which dies in the final step of the filtration. The zeroth Betti number equals the number of connected components. The first Betti number equals the number of loops **(b)**. We use a barcode to visualise the birth and death of the Betti numbers **(c)**.

Persistent homology of weighted networks

- To reduce computational efforts, in persistence homology, networks are represented as clique complex. Remember, this complex fill in all triangles and tetrahedrons which will not be counted as loops. As a result, the only loops detected have four or more vertices with no diagonal connections.
- Filling triangles does not prevent us from knowing the number of connected components because we can always analyze only the 1-skeleton which has no loop.

Collaboration Networks

- Persistence homology is applied to 4 collaboration networks of scientists.
- The vertices= author of the papers
- Edges= co-authors on at least one paper
- Weight= 1\(n-1) with n= number of authors of the paper
- As a result, different communities are strongly connected and weak ties are necessarily part of communities.



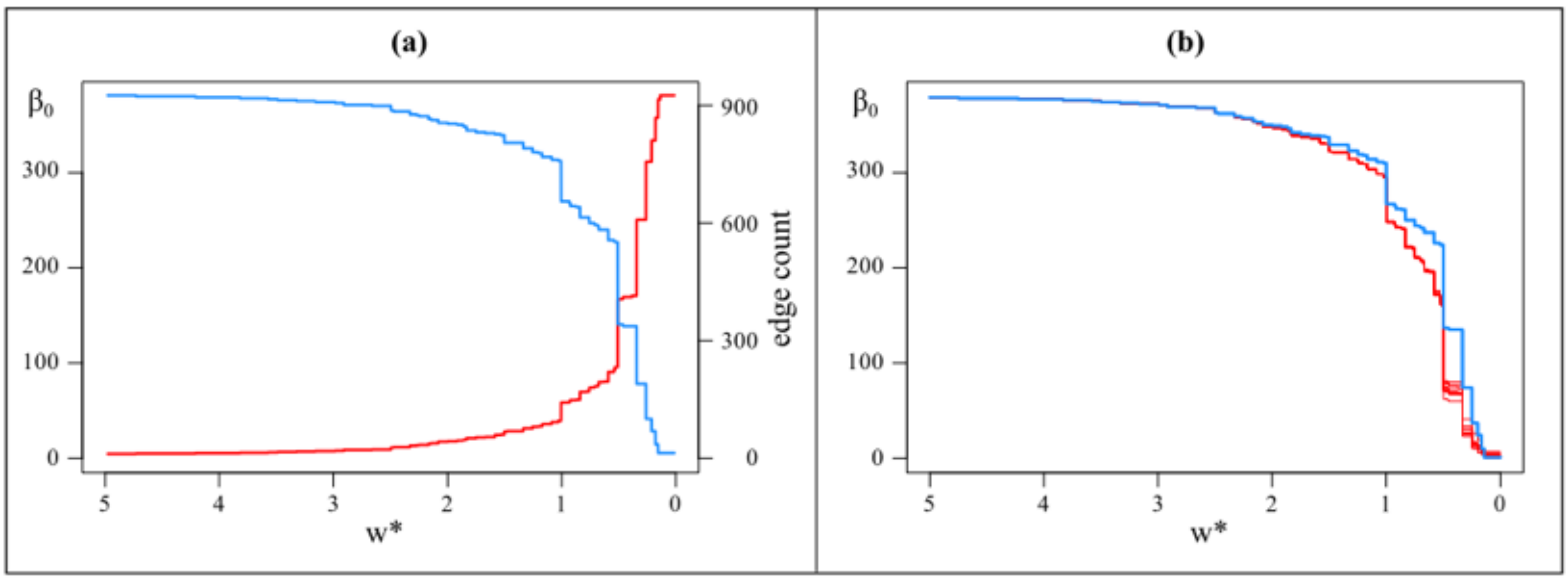


Figure 3: On the left (**a**) we plot the zeroth Betti number against the threshold w^* in blue. The total number of edges present at each stage of the filtration is plotted in red. On the right (**b**) we again plot the zeroth Betti number in blue. There are ten red plots, each corresponding to the sequence of zeroth Betti numbers of a random graphs with the same number of vertices, edges and the same weights.

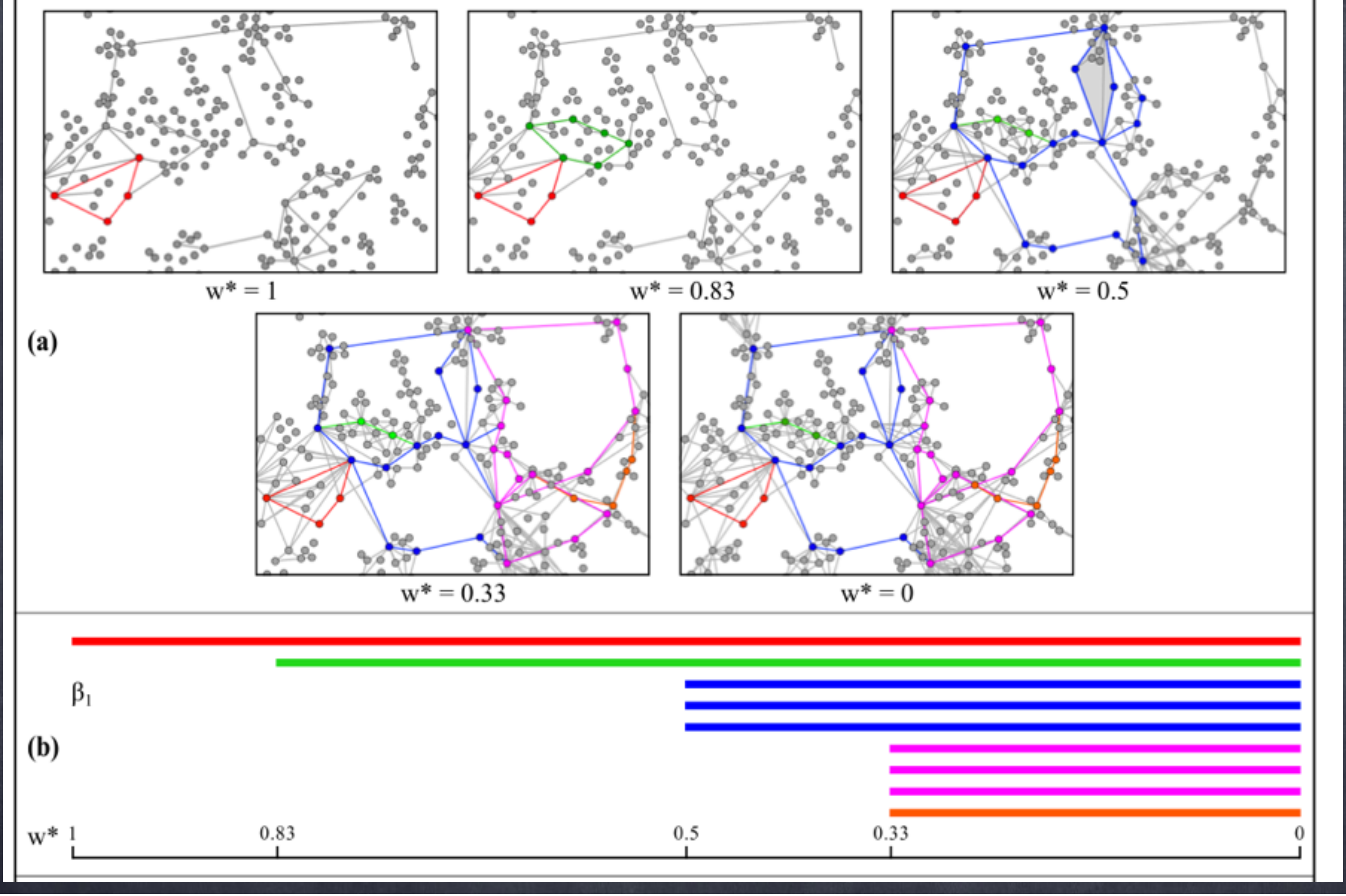


Figure 4: We only show the central part, see Figure 2, of the network of 379 network scientists since all loops occur here **(a)**. We find the first relatively small loop between four scientists for appearing for threshold weight 1. As we decrease the threshold weight more loops appear. Notice how we have shaded two triangles for $w^* = 0.5$, this is to indicate that there is no loop there, there are three new blue loops added at this stage. We notice that for smaller threshold values we find larger loops. In **(b)** we show the barcode for the first Betti numbers of this filtration. In **(c)** we enlist the length of the loops that appear at each filtration value.

For Random Network, we obtain higher Betti numbers because there are less clustering and thus fewer triangles that are filled in and more loops with more than three edges.

Conclusion

- By applying persistent homology to four collaboration networks of scientists we have shown that it gives us interesting information about the structure of the networks.
- We found that due to the construction of collaboration networks, weak ties form cliques and strong ties act as local bridges between those cliques. This is contrary to what has been described in other social networks.
- We were also able to use persistent homology to distinguish these collaboration networks from random networks. Using the one and two dimensional Betti numbers of the network we did not need to take the weights into account.