Matt Howa

Modularity Research Update

4/4/2017

1. Output Diagrams
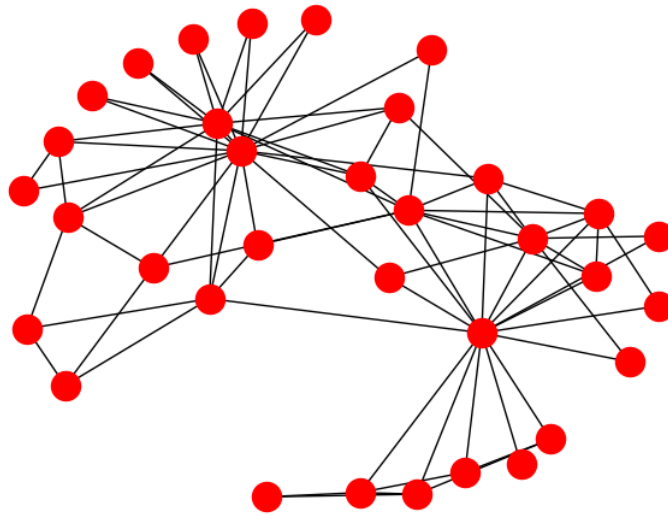

Figure 1: *Original Karate Club Network*
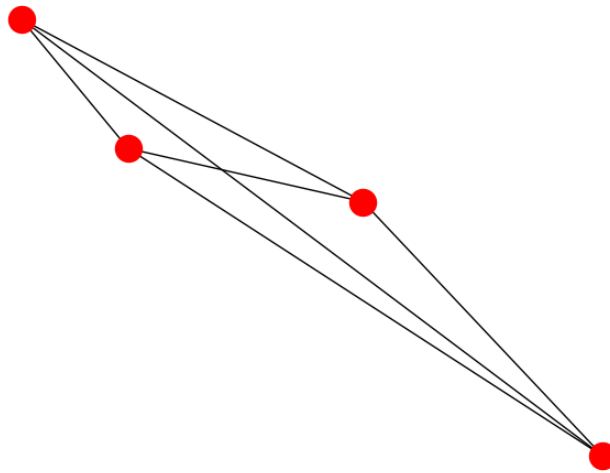

Figure 2*: New Karate Club Network after running Clustering using Modularity.*
Modularity Score: 0.375986193294
*The edge lengths have been calculated based on the number of original connecting edges from
one cluster to the next and weighted accordingly.*

a. From the two graphs, I ran shortest path algorithms to create a pairwise distance matrix.
b. Then I ran both of the matrices through Dipha to compute their Persistent Homology and output a Persistent Diagram for each

Here are the P X 3 Matrixes for comparing Homology

ORIGINAL GRAPH

	[, 1]	[, 2]	[, 3]
[1,]	1	0	2.5
[2,]	0	0	0.5
[3,]	0	0	0.5
[4,]	0	0	0.5
[5,]	0	0	0.5
[6,]	0	0	0.5
[7,]	0	0	0.5
[8,]	0	0	0.5
[9,]	0	0	0.5
[10,]	0	0	0.5
[11,]	0	0	0.5
[12,]	0	0	0.5
[13,]	0	0	0.5
[14,]	0	0	0.5
[15,]	0	0	0.5
[16,]	0	0	0.5
[17,]	0	0	0.5
[18,]	0	0	0.5
[19,]	0	0	0.5
[20,]	0	0	0.5
[21,]	0	0	0.5
[22,]	0	0	0.5
[23,]	0	0	0.5
[24,]	0	0	0.5
[25,]	0	0	0.5
[26,]	0	0	0.5
[27,]	0	0	0.5
[28,]	0	0	0.5
[29,]	0	0	0.5
[30,]	0	0	0.5
[31,]	0	0	0.5
[32,]	0	0	0.5
[33,]	0	0	0.5
[34,]	0	0	0.5

CONDENSED GRAPH

	[, 1]	[, 2]	[, 3]
[1,]	1	0	1.0
[2,]	0	0	0.5
[3,]	0	0	0.5
[4,]	0	0	0.5

After computing bottleneck distance in TDA the output was 0 distance. Im afraid this is wrong and I need to look into it. I believe the answer is with the dipha program
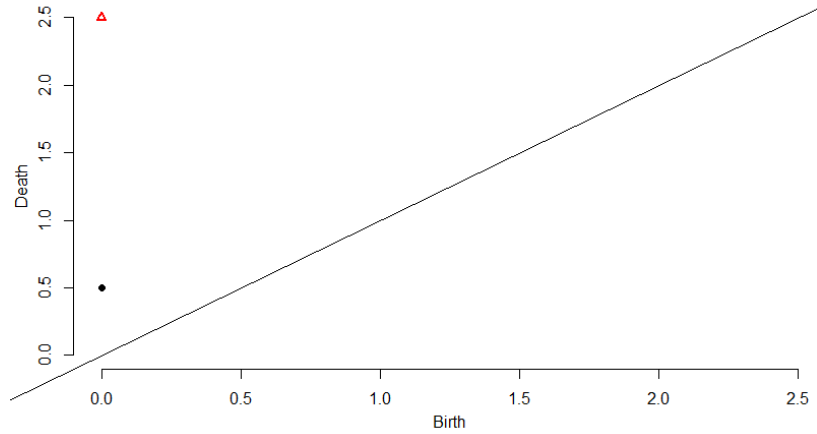
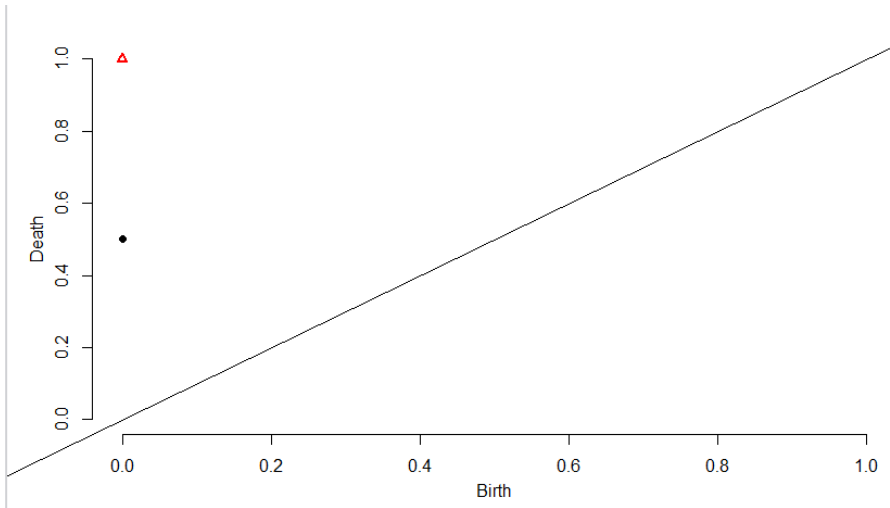***Figure 3:*** *Rips-Diagram of Original Graph*

Figure 4*: Rips-Diagram of Condensed Graph*

At this point it is clear my computation of persistent homology through dipha is not working or is configured incorrectly. My new goal is to look into these settings

2. Dendrogram Explanation

A dendrogram is a tree representation of hierarchical clustering. When we cluster, nodes using modularity, we create a dendrogram which represents an ordering to which we merge nodes. Choosing where to partition a graph becomes easy after we create the dendrogram because we can pick the best communities based on different levels of the dendrogram. From the dendrogram we can choose the best partition of communities and from it create a condensed graph. I've implemented this specifically by taking the best partition or level where the

communities modularity is the highest. From here I merge the communities into a single node and reweight the edges to reflect how many edges originally interconnected the communities. Finally using the new edge weights, I compute the shortest paths between the new (super) nodes. This new pairwise distance matrix is then used to compute persistent homology.