Homework

Group 4

Wojciech Bogucki
Bogdan Jastrzębski
Karol Pysiak
Tomasz Makowski

Task 4

MCODE

Firstly, we have run MCODE algorithm with default settings. As we can see below, red nodes corresponds to nodes without cluster, green corresponds to clustered nodes and purple nodes are seeds. Not many nodes were assigned to the cluster. Only one cluster was obtained.

On the right the same algorithm but with “fluff” parameter on. As we can see more nodes were assigned to cluster but still only one cluster appears.

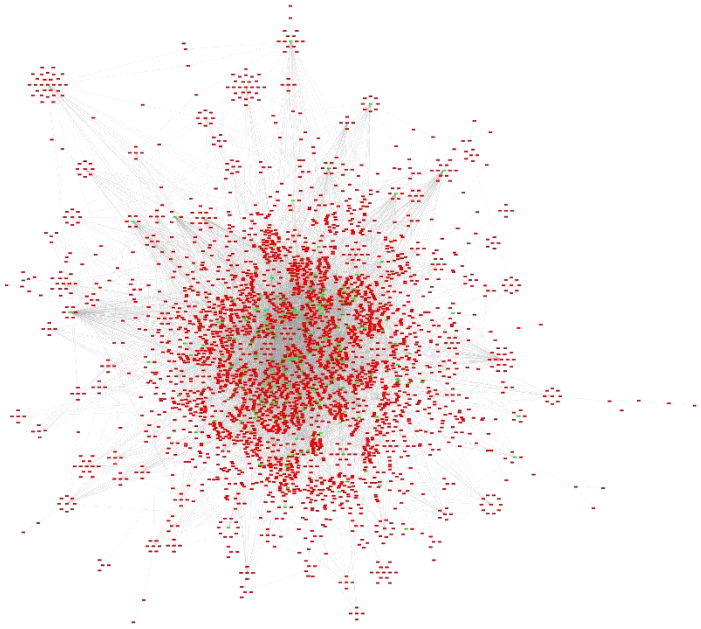


Figure 1 Clustering by default MCODE algorithm

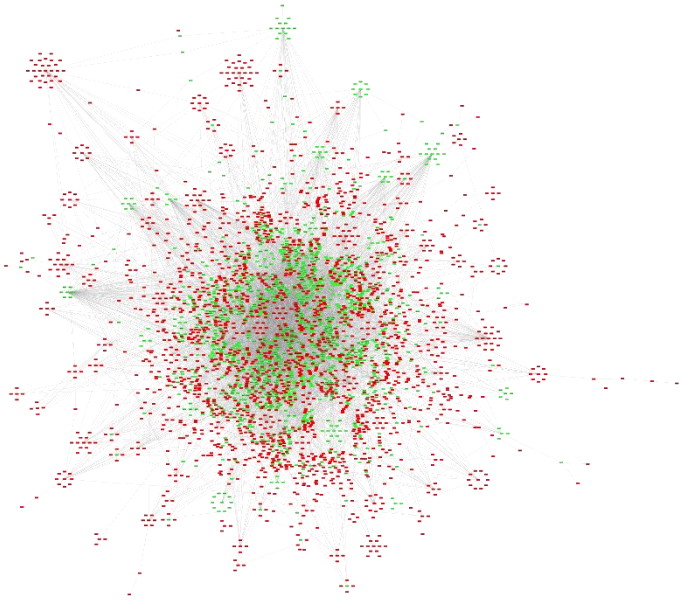
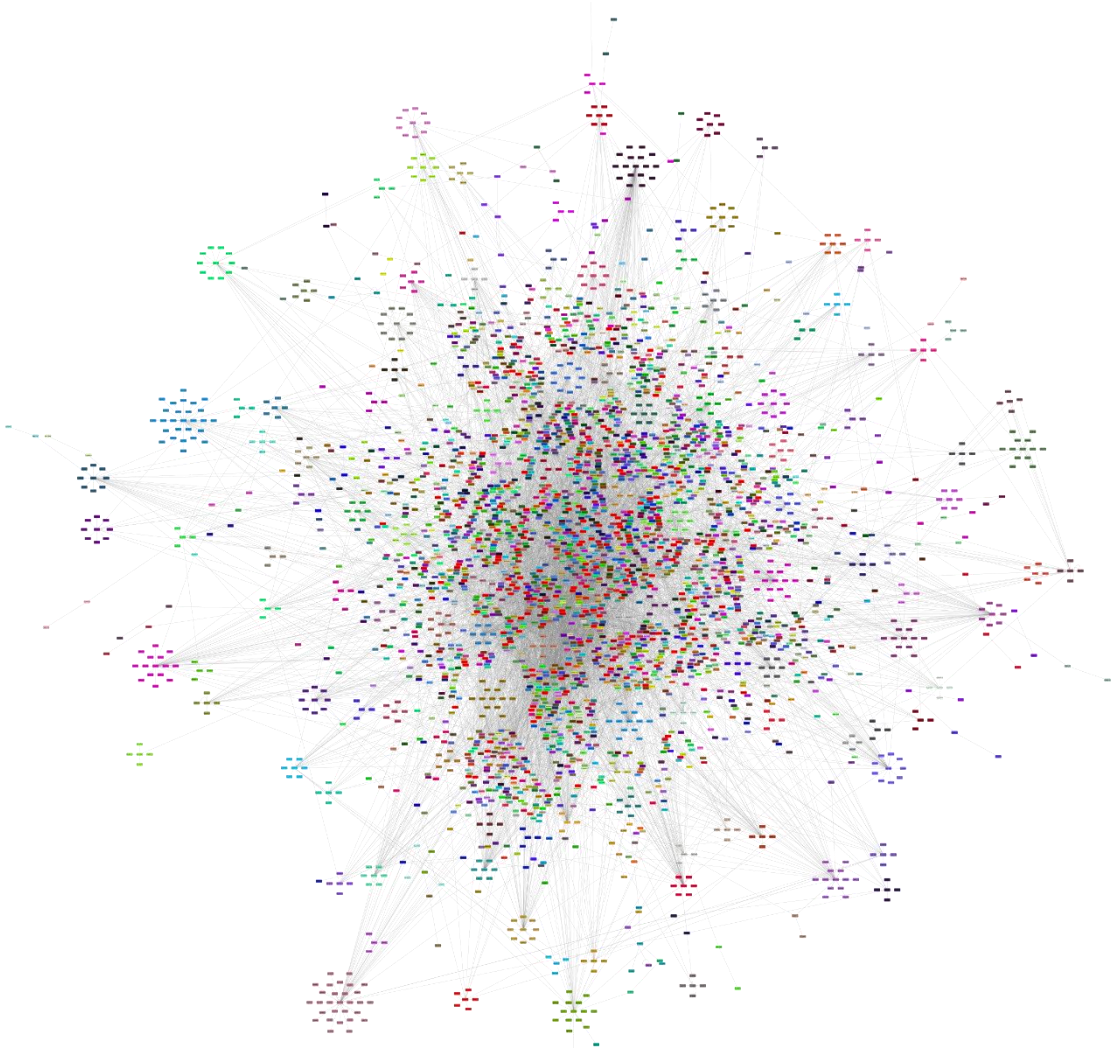


Figure 2 Clustering by MCODE algorithm with fluff option

MCL

Next we have run MCL cluster algorithm from clustermaker2 package with default settings. As we can see below, much more nodes were clustered by this algorithm. We obtained 852 clusters.



Change of granularity parameter resulted in 724 clusters and more node assigned to some cluster.

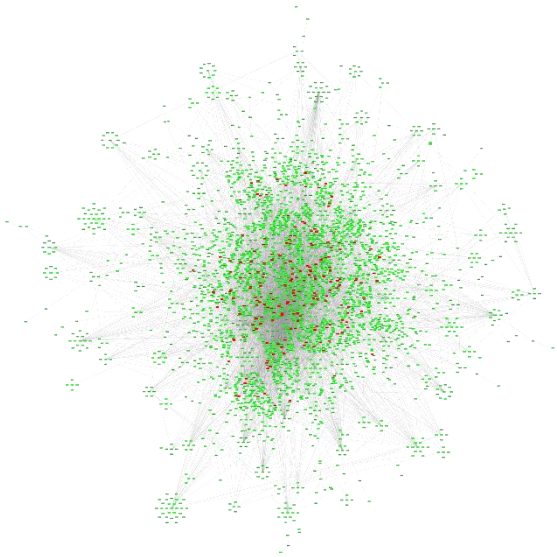


Figure 3 Granularity parameter 2,5 (default)

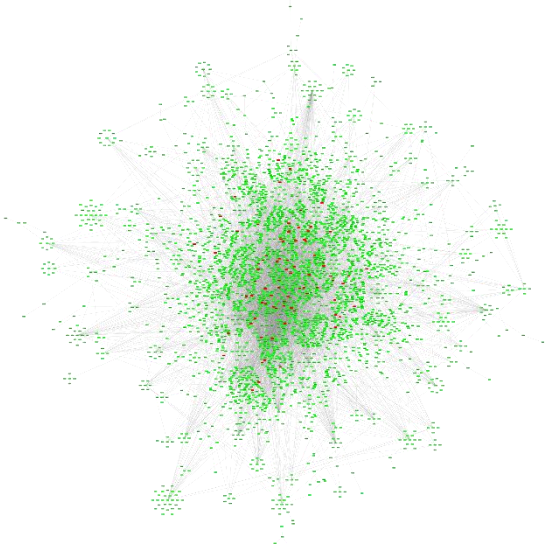


Figure 4 Granularity parameter 2,0. There is less red nodes (unclustered) visible in the network

Changing array source to EdgeBetweenness (edge cutoff resulted in 198 clusters. All nodes were assigned to clusters.

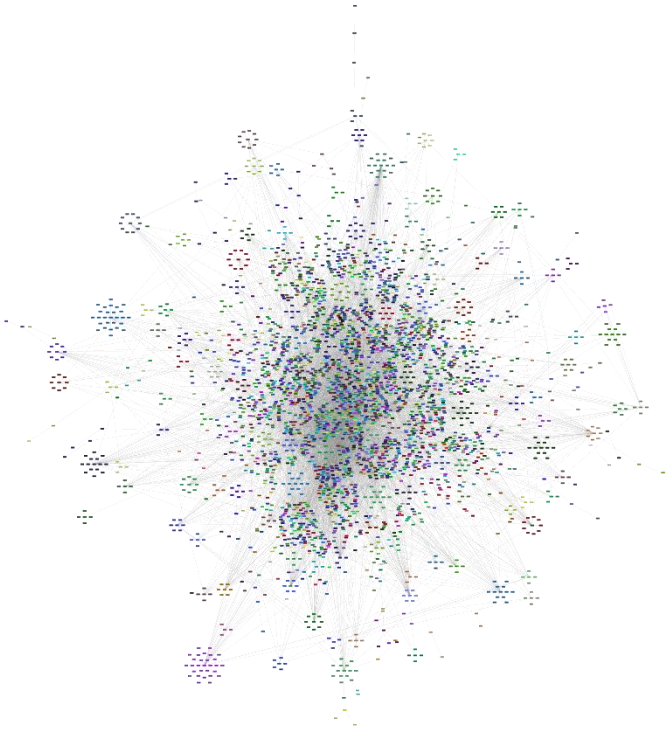


Figure 5 Drosophila network divided into clusters by MCL algorithm

Conclusions

MCODE algorithm with default setting has problem with clustering drosophila network – many nodes without cluster and the rest in one cluster. MCL algorithm works better and find a lot of cluster which can be reduce by setting parameters like granularity and array source.